MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 111 No. 4 1951

Published and Sold by the
ROYAL ASTRONOMICAL SOCIETY
BURLINGTON HOUSE
LONDON, W. 1

Price Nine Shillings

NOTICE TO AUTHORS

1. *Communications.*—Papers must be communicated to the Society by a Fellow. They should be accompanied by a summary at the *beginning* of the paper conveying briefly the content of the paper, and drawing attention to important new information and to the main conclusions. The summary should be intelligible in itself, without reference to the paper, to a reader with some knowledge of the subject; it should not normally exceed 200 words in length. **Authors are requested to submit MSS. in duplicate. These should be typed using double spacing and leaving a margin of not less than one inch on the left-hand side. Corrections to the MSS. should be made in the text and not in the margin.** Unless a paper reaches the Secretaries more than seven days before a Council meeting it will not normally be considered at that meeting. By Council decision, MSS. of accepted papers are retained by the Society for one year after publication; unless their return is then requested by the author, they are destroyed.

2. *Presentation.*—Authors are allowed considerable latitude, but they are requested to follow the general style and arrangement of *Monthly Notices*. References to literature should be given in the standard form, including a date, for printing either as footnotes or in a numbered list at the end of the paper. Each reference should give the **name and initials** of the author cited, irrespective of the occurrence of the name in the text (some latitude being permissible, however, in the case of an author referring to his own work). The following examples indicate the style of reference appropriate for a paper and a book, respectively :—

A. Corlin, *Zeits. f. Astrophys.*, **15**, 239, 1938.
A. S. Eddington, *Internal Constitution of the Stars*, Cambridge, p. 182, Table 24, 1926.

3. *Notation.*—Authors should conform closely to the recommendations of Commission 3 of the International Astronomical Union (*Trans. I.A.U.*, Vol. VI, p. 345, 1938). Council has decided to adopt the I.A.U. 4-letter abbreviations for constellations where contraction is desirable (Vol. IV, p. 221, 1932).

4. *Diagrams.*—**These should be drawn about twice the size required in print and prepared for direct photographic reproduction except for the lettering, which should be inserted in pencil. Legends should be given in the manuscript indicating where in the text the figure should appear.** Blocks are retained by the Society for 10 years; unless the author requires them before the end of this period they are then destroyed.

5. *Tables.*—**These should be arranged so that they can be printed upright on the page.**

6. *Proofs.*—Costs of alterations exceeding 5 per cent of composition must be borne by the author. Fellows are warned that such costs have risen sharply in recent years, and it is in their own and the Society's interests to seek the maximum conciseness and simplification of symbols and equations consistent with clarity.

7. *Revised Manuscripts.*—When papers are submitted in revised form it is especially requested that they be accompanied by the original MS.

Reading of Papers at Meetings

8. When submitting papers authors are requested to indicate whether they will be willing and able to read the paper at the next or some subsequent meeting, and approximately how long they would like to be allotted for speaking.

9. Postcards giving the programme of each meeting are issued some days before the meeting concerned. Fellows wishing to receive such cards whether for Ordinary Meetings or for the Geophysical Discussions or both should notify the Assistant Secretary.

MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 111 No. 4

MEETING OF 1951 APRIL 13

Professor H. Dingle, President, in the Chair

The election by the Council of the following Fellows was duly confirmed:—

Cecil Gilbert, 1 North House, Longbenton, Newcastle-on-Tyne (proposed by K. Mitchell);

Bernardo Grossling-Freudenburg, M.Sc., 26 Emperor's Gate, London, S.W.7 (proposed by A. T. Price);

William Halliday, B.Eng., 3 Hollycroft Avenue, London, N.W.3 (proposed by P. F. Everitt);

Avril Barbara Hart, University Observatory, Oxford (proposed by H. H. Plaskett);

John Warren Loveridge, Middleheath, 11 West Heath Road, London, N.W.3 (proposed by A. Hunter); and

George Tait, 16 Avenue Road, St. Albans, Hertfordshire (proposed by R. W. B. Pearse).

The election by the Council of the following Junior Members was duly confirmed:—

Bernard Ephraim Julius Pagel, The Observatories, Madingley Road, Cambridge (proposed by R. H. Garstang); and

Dennis William Scaima, The Cavendish Laboratory, Free School Lane, Cambridge (proposed by R. H. Garstang).

One hundred and fourteen presents were announced as having been received since the last meeting, including:—

J. B. Sidgwick, *Introducing Astronomy* (presented by the author); and

W. M. Smart, *The Origin of the Earth* (presented by the author).

The President announced that the Council had elected the following Associates of the Society:—

Dr Gerard P. Kuiper, Yerkes Observatory, Williams Bay, Wisconsin, U.S.A.; and

Professor Polydore Swings, Institut d'Astrophysique, Université de Liège, Sclessin, Belgium.

Professor Antonie Pannekoek delivered the George Darwin Lecture (see p. 347), taking as his subject "The Origin of Astronomy".

After the lecture, the President delivered a short address, and presented the Gold Medal of the Society to Professor Pannekoek, to whom it had been awarded for his work on astrophysics and on the structure of the galactic system. The President said:—

"I suppose what strikes us first about the subject of this lecture which we have just heard is its remoteness from the kind of paper which we have usually heard at our meetings. Yet Professor Pannekoek has dealt with the actual foundations on which our more advanced work is built. He has mentioned that in Babylonian times they would be prepared to bet five to one on the success of an eclipse prediction. Anyone who did the same to-day with some of the predictions of modern theories would be regarded as a gambler! The debt which the Greeks owed to Babylonian observations forms an interesting study, and papers on this subject are so few that we are fortunate to have had one at this meeting.

"Professor Pannekoek, it is now my pleasure to hand to you the Gold Medal of the Society, which is the highest honour the Society has it in its power to bestow. Will you accept it with our thanks for your past work and our best wishes for further enjoyment in the many fields of astronomy in which you have gathered so rich a harvest."

MEETING OF 1951 MAY 11

Professor H. Dingle, President, in the Chair

The election by the Council of the following Fellows was duly confirmed:—

*Richard Myer Baum, Stocks Lane Post Office, 1 Dee Banks, Boughton, Chester (proposed by P. A. Moore);

Ronald Winston Cable, 77 Holland Street, Wembley, Perth, Western Australia (proposed by R. W. Boggis);

Selwyn Rhys Davies, B.Sc., 12 Brohanddgar, Felinfoel, Llanelly, South Wales (proposed by B. E. Featherstone);

Thomas Charles Mansfield, M.A., 7 Knole Road, Sevenoaks, Kent (proposed by H. Spencer Jones);

Tadeusz Boleslaw Slebarski, M.A., University Observatory, St. Andrews, Scotland (proposed by E. Finlay-Freundlich);

Stanislaus Alexander Vincenz, B.Sc., A.R.C.S., D.I.C., Imperial College of Science, South Kensington, S.W.7 (proposed by J. McG. Bruckshaw); and

Andrew Young, B.Sc., Department of Applied Mathematics, The University, Liverpool (proposed by E. Finlay-Freundlich).

One hundred and ten presents were announced as having been received since the last meeting, including:—

J. W. Mitchell, *Fundamental Mechanisms of Photographic Sensitivity* (presented by the editor).

* Transferred from Junior Membership.

THE ORIGIN OF ASTRONOMY

George Darwin Lecture, delivered by Professor Antonie Pannekoek on 1951 April 13

The origin of astronomy, as of all science, lies far back in prehistoric times. Primitive man in his struggle for life could not do without a certain knowledge of the surrounding natural phenomena; the better he was acquainted with them, the better his life was secured. Thus, when he reached the cultural stage which ethnology calls the higher stage of barbarism, a first, merely practical, knowledge of physics, chemistry and biology was contained in his techniques, his agriculture and his cattle raising. At this stage some knowledge of celestial phenomena must also have been present, just as to-day we find it with primitive tribes. Whence did it come? Sometimes the opinion has been expressed that it was the beauty of the starry heavens, the mysterious silent course of its luminaries with their continuous change of aspect, that attracted the attention of man and awakened the desire to find out their cause. This opinion is clearly a reflex of modern conditions, where physics and chemistry are closely tied up with economic life, whereas astronomy stands outside the direct necessities of life and is studied for its own sake, for the beauty of its object and the wide scope of its problems. In opposition to this Schiaparelli (1) rightly argued that the first study of astronomy was imposed upon man by the desire to satisfy the necessities and comforts of his life. Indeed, economic necessities were at the root of his astronomical as well as of his physical and chemical knowledge. The heavens above him belonged as an essential part to the surrounding world which he had to know.

So, to understand the origin of astronomy, we have to see what practical necessities directed the attention of primitive man towards the celestial phenomena. There is first the need for orientation or direction finding. When commerce had developed out of nomadic cattle-herding or coastal fishery, the caravans in the desert and the ships at sea needed the celestial luminaries to guide them. During daytime the Sun and at night the stars directed their travels. Thus the Arabs from olden times were acquainted with a number of stars, especially with the 27 groups which the Moon successively visits in its monthly course. In Homer's Odyssey the goddess Kalypso instructs Odysseus in sailing home to keep the Bear at his left hand. The Polynesian island dwellers in the Pacific, before their contact with Europeans, were well acquainted with the stars; as skilled mariners they steered their vessels by means of the rising and setting points of different stars, as a kind of celestial compass; they had schools for navigation where, by means of globes, young people were instructed in astronomy.

More important still and more universal than for this orientation in space was the use of the heavenly bodies for what may be called the orientation in time, the time reckoning, the calendar. All work of man, at least before the industrial revolution, was bound up with the alternation of the seasons. The natural phenomena on which his activities depend were regulated by the period of the solar year. In Northern and moderate climates they depended on the alternation of a dead winter and a life-abounding summer, in Southern and tropical countries mostly on the alternating of a dry and a wet season. As hunters or fishermen they had to follow the migrations of the animals; as herdsmen or farmers they had to regulate their work according to the seasonal life-periods of animals and plants.

Thus, in the planning of his work, in order not to be misguided by irregularities of the weather, man had to have some independent means for counting the days and recognizing the right time for sowing and harvesting. Such means were provided by the Sun itself, and by other celestial phenomena.

The phenomena connected with the yearly variation of the Sun's declination are *eo ipso* exact indications of the seasons, and were recognized and used as such. Many primitive peoples to-day are found to use the solar phenomena as signs of the time of year. Dayak tribes measure the shadow of a vertical stick held at arm's length to know the month (**2**). The Eskimos and the Zuni Indians are reported to observe the solar solstices by watching the moments of extreme rising and setting at points of the horizon which are marked by stones (**3**); such moments are important dates celebrated by religious rites.

Another important characteristic of the season is given by the stars, most manifestly by their heliacal rising, in the morning just before sunrise. The Australian aborigines at Torres Strait for the beginning of sowing await the appearance of a bright star called Kek, probably Canopus or Achernar, for which a close watch is kept (**4**). Analogous practices are mentioned for other primitive peoples. That the same is true of prehistoric man is indicated by fragments of the oldest historical testimonies. Everybody knows how in ancient Egypt the heliacal rising of Sirius in the month of June was taken to announce the flooding of the Nile. From ancient Greece, Hesiod's poem *Works and Days* mentions for the different rural activities what phenomena of the stars are connected with them; e. g. "When the Atlas-born Pleiades rise, then begin mowing, but ploughing when they disappear" (**5**).

But more than the Sun and the stars, the Moon occupied the chief place in time-reckoning. Its regular changes of phase offered the most practical period for counting the days. Everywhere, with all peoples—even those who afterwards adopted, as we do, pure solar reckoning—the Moon's synodic period of $29\frac{1}{2}$ days was originally the basis of the calendar. The month begins with the first appearance of the new sickle Moon as a slender arc in the Western evening sky and as a consequence the day begins with the evening. The Moon's gradual waxing and waning impressed man as the life-history of a living being, ever again new-born. In the sacred books of many peoples it is said that the Moon has been created for the purpose of measuring the time. The Moon was the oldest and most venerable of the gods; its first appearance was attentively watched and hailed, and still more its fullness, when it dispels the darkness of the night, was celebrated with religious ceremonies.

Thus a certain knowledge of celestial phenomena, as a part of surrounding nature, is present in prehistoric man, involved with all his other technical knowledge in the practical activities of sustaining his life. Often it was a speciality in the hands of old people or priests, who performed the small functions of spiritual leadership in these communities. We cannot truly call it astronomical science, any more than his techniques of making tools and preparing food is called physical or chemical science. Science in the real sense of the word could only arise when man entered into a higher stage of development, the stage of civilization.

The transition from the prehistoric stage of barbarism to civilization is marked by the invention of script. This means more than that from now on written history supplants the oral tradition of legends. The importance of script is that it gives visible form to the conceptions that before had only a spiritual

existence in our mind. The words of the spoken language expressing these conceptions had already been in use for countless generations, as a means of mutual intercourse and incentive to action, but only as part of instinctive practical life; once the fleeting sound had passed, nothing material remained. Through script, however, conceptions acquire an existence of their own, visible and permanent; they can be handled and stored, compared and connected with one another. Now science as a system of generalized knowledge, of relations between general notions and abstract conceptions, becomes possible. Science is theory; it deals with things in general, with the abstractions formed by our mind out of the phenomena. Thus the rise of science as theoretical knowledge must be seen as part of the larger historical process, the rise of civilization.

The origin of civilization has to be sought in the realm of social developments. "War and Class", writes the English historian Arnold Toynbee, "are the two congenital diseases of civilization" (**6**). Now war was inherent in the preceding stages too, during savagery and barbarism. But "Class", i.e. the separation of a ruling class from the working masses, has been the social characteristic of civilization only, in all its forms. Its role in the development of astronomical science must now be traced. The countries to which we have therefore to turn our attention are Babylonia and Greece.

The first civilizations arose about 5000 years ago in the fertile plains of Egypt, Mesopotamia, India and China. Here the silt deposits of the great rivers, Nile, Euphrates, Hoang-ho, offered a good soil of marvellous fertility for a numerous population, but continual care was needed to regulate the water supply by dykes and canals. At regular or irregular moments the river broke through the dykes, flooded the land and deposited new layers of silt. This care could not be left to the separate villages or districts with their often opposed interests and so a strong central power was necessary to provide for the common interest. The multitude of small communities with their own chieftains and deities consolidated into large monarchies. This was rendered possible because the fertility of the soil procured abundant surplus produce for the sustenance of a separate class of ruling officials.

Moreover a strong state power was needed for defence against the poor warlike tribes which in the surrounding mountains and deserts could find a meagre living only. They made it their business to attack and plunder their prosperous neighbours. So a division of labour was necessary; an army of soldiers had to protect the farmers and in time became a ruling group, with their chief as king taking into his hands the judicial and administrative powers. Sometimes the plundering tribes became conquerors and settled as a ruling aristocracy among the subjected farmers; theirs was now the task of staving off new aggressors. The result was the same in both cases. Beside the military power stood the civil officials. With the separate townships organized into a large empire their local deities too were organized into a pantheon and their local priests into a powerful hierarchy, the spiritual leaders of society.

With the appearance of these ruling classes new needs arose. Their greater wealth expressed itself in luxury, in the development of refined arts, in spiritual culture. The conditions were now present for the introduction and spreading of script. The ruling monarchs wanted to glorify their deeds, their victories in war and their temple-building in peace-time. Missives with reports and instructions were interchanged between the king and his officials and governors, and his decrees were fixed as laws by writing them down. Writing as a special

capacity of priests and scribes was also put at the service of private individuals for their business contracts. Thousands of such contracts as well as other documents, written on clay in cuneiform characters, have been dug up from temple ruins of the Babylonian towns and are preserved in our museums.

What was the effect of the new conditions upon the rise of astronomy? A special group of officials had now to pay attention to the celestial phenomena. The regulation of time was one of the chief tasks of the priesthood; from their staged towers they had to observe the first appearance of the new sickle Moon and to announce the new month. Then the Moon reckoning had to be adapted to the seasons of the year. Because 12 periods of the Moon are 11 days shorter than the solar year, every 3 or 2 years a 13th month must be intercalated, not for any theoretical or scientific reason, but in order that the offering of the first crops to the gods may take place at the prescribed time, at the full Moon ceremonies of the first month Nisannu. In such primitive societies the important agricultural activities were at the same time religious festivals; economics and worship, politics and religion formed an indissoluble unity. The service of the gods was bound to strict ritual; the calendar, the time regulation of worship, was a sacred duty of the priests, the source of their prestige and spiritual power. Originally the intercalation of a 13th month took place according to practical necessity; when at the end of the 12th month it appeared that ripeness of the crops and the beginning of harvesting could not be expected in the next fortnight, the 12th month had to be repeated. The inscriptions show that about 2000 B.C. the intercalations took place at irregular intervals, and hence were probably determined by practice.

This may have caused annoyance, but a better way presented itself. When looking out for the new Moon the priests must have perceived that in every succeeding month the surrounding stars had changed, gradually moving farther West till they disappeared, being replaced by more easterly constellations. At the same time new stars rose in the morning twilight. So the phenomena of the stars indicated the seasons and could afford a more regular and exact means for intercalation of the additional months. There are many texts where the names of constellations are connected with the names of months, indicating their use for calendar purposes. An easy rule is given in a text published by George Smith: "When at the first of Nisannu the Moon and star Mulmul (the Pleiades) stand together the year is common; when at the third day of Nisannu the Moon and the star Mulmul stand together the year will be full" (7). The latter case means that the first month falls so early in the seasonal year that addition of a 13th month becomes necessary. As a consequence we find that long before 1000 B.C. the Babylonians were acquainted with a considerable number of constellations, by names in inscriptions and by pictures on boundary stones. They are to a great extent identical with our own, evidently transmitted by the Greeks. That in these same centuries also the brightest of the planets, with its alternation as evening and morning star, was perceived, is evident from the famous so-called Nindar-anna tablet from about 1600 B.C., where its phenomena are described, and their significance.

For in the course of time it was no longer only, or chiefly, calendar needs that determined man's interest in the stars. Ever more astrology, the doctrine of the significance of the celestial phenomena for the events on Earth, occupies his mind. Now what happens in the sky is followed with keener interest as signs of the destiny allotted to people and states. Belief in signs and presages is as universal in primitive man as his belief in invisible spiritual powers surrounding

him and influencing his work and life. To win their favour, to find out their intentions, to avert or appease their hostility, by magic charms and incantations is an important part of his daily practice. Most of these spirits have their abode in the heavens; and especially in these Mesopotamian plains, with their resplendent sky, the stars impressed man's mind by their radiant brightness. Thus in early times the conception of a close connection between the stars and man's fortunes had already risen in the mind of the Babylonian priests. Old Sumerian inscriptions speak of the "favourable star" for temple building. This belief was alive through all the centuries when Babylon, as a rich commercial town and the capital of a large empire, had become a centre of culture for the surrounding world. But it intensified to an all-pervading fervour when in the centuries after 1000 B.C. Assyria rose and became the most powerful state in the Near East, extending its dominance not only over Babylon and all Mesopotamia but also over Syria, Palestine and even Egypt. In the vicissitudes of world politics at the time and in the variable fortune of war the need was strongly felt for aids in foreseeing the future. So the astrologers in direct service of the king had to give their omens before every great enterprise, and from all the great sanctuaries the priests had to send regularly their reports on what they had seen in the heavens and the meaning thereof. From the ruins of the library of King Assurbanipal ten thousands of more or less damaged sherds with cuneiform inscriptions have been dug up, many of them containing astrological reports. These enable us to obtain a clear picture of the character and scope of their astronomical observations.

The chief objects observed were the Moon and the planets, because by their variable aspect and irregular wanderings they offered the greatest variety of phenomena suitable for astrological interpretation. The Moon was now observed with more attention in all its phases; especially were its risings and settings observed in the middle of the month, about full Moon. When the Moon is full on the night of the 14th, the normal time, it was a lucky omen; when full Moon happened on the night of the 13th, 15th or 16th, it was abnormal, hence a bad omen. Here astrology and calendar were merged; deviation in the calendar was considered an unlucky sign and had to be restored at the end of the month. The eclipses, of course, were important phenomena, mostly unfavourable; but as to the different sides of the Moon that were successively obscured and the different hours of the night, they had significance for different countries, for Akkad, the southern Babylonia proper, for Elam in the East, for Amurru in the West. So the details were often so well noted that nine centuries afterwards Ptolemy could use them in the construction of his lunar theory.

But the most important effect of astrology was that now besides the Moon the planets strongly captured the attention of man. They were seen wandering between the stars in the most intricate and unexpected ways, as living beings roaming through the starry landscape. They are the stars of the great gods who rule the world and manifest themselves in these bright luminaries. Their appearances and disappearances, their motions through the constellations of the zodiac, their stationary points and retrogradings, their meeting with one another and with bright stars, presented an almost endless variety of phenomena. Although from earlier times we have already a text with omens from Venus, it is in these later centuries of Assyrian power that a high tide of astrology provides us with an abundance of reports containing omens from all the five planets. There we read (**8**): "Mercury went back as far as the Pleiades"; "Jupiter enters Cancer"; "Venus appears in the East"; "Mars is very bright"; "Jupiter appears in the

region of Orion"; "Mars stands in Scorpio, turns and goes forth with diminished brilliancy"; "Saturn has appeared in the Lion"; "Mars approached Jupiter"; and so on. There is not a trace of scientific interest in these texts; the mind of the reporters is entirely occupied by the omens: when such or such happens, "it is lucky for the king, my lord"; or, "copious floods will come"; "there will be devastation"; "the crops will be diminished"; "the king will be besieged"; "the enemy will be slain"; "there will be raging of lions and wolves"; "the gods intend Akkad for happiness"; and so on. Yet, with all those observations, these reports represent a considerable astronomical activity. For the first time in history a large number of data on the planets had been collected; it implies a detailed knowledge of facts about their motion.

This does not mean that we have here the beginning of a science of astronomy. Mere facts do not form a science. Facts are the basis of science. Science is the systematization of facts into general conceptions and rules. Now it might be imagined that these ancient observers, by comparing their observations with old records and looking for regularities, must gradually have detected them. But we must beware of regarding these Babylonian priests as astronomers like ourselves, animated by the same spirit of scientific purpose and scientific research. Such a state of mind certainly was entirely absent then. It is the very problem of the origin of science to discover how it originally came about that, without knowing about scientific aims, people were able to build up scientific theory.

The circumstance that made this possible for astronomy was the occurrence of extremely simple and striking periodicities in the celestial phenomena. What looked irregular on occasional and superficial observing revealed its regularity in a continuous abundance of data. Regularities were not sought for; but regularities imposed themselves, without giving surprise. They aroused certain expectations. Expectation is the first unconscious form of generalized knowledge, like all technical knowledge in daily life growing out of practical experience. Then gradually the expectation develops into prediction, an indication that the rule, the regularity, has entered consciousness. In the celestial phenomena the regularities appear as fixed periods, after which the same aspects return. Knowledge of the periods was the first form of astronomical theory.

Did the astrologers in Assyrian times already possess such knowledge? There are some texts indicating that they did. In one of them (9) we read: "To the king, my lord, I sent: an eclipse will take place. Now it has not passed, it has taken place. In its happening the eclipse portends peace for the king, my lord." Here it is stated that a former prediction has been confirmed. The prediction of a lunar eclipse indeed is very simple, because 5 or 6 such eclipses follow one another always regularly with 6 months' interval; and when this series has ended a new series starts 11 or 17 months later. Hence, if an eclipse has been observed we may bet five to one that 6 months later another will take place. There is a difficulty, however; on the average half of them will be invisible because they occur in the daytime, when the full Moon is below the horizon. This irregularity must have hampered the easy discovery of the continuous series by the Babylonian priests. But their own regular observations of the Moon's risings and settings in the midst of the month must have shown them a way towards understanding such gaps. When full Moon happens in the night, then in the morning the Moon will set only after sunrise; Sun and Moon, god Shamash and god Sin, are seen at the same time, one East, the other West. If however, the Moon does not wait for the Sun but sets before sunrise, full

Moon has still to come; and if in such a case an eclipse was expected, it will remain invisible. That the Babylonian priests were aware of these conditions is shown by a published text, probably from the 7th century B.C.: "The eclipse passes, it does not take place. If the king should ask: what omens hast thou seen? —the gods have not been seen one with another" (**10**). Here clearly the eclipse was predicted, and now the reason is given why it must remain invisible.

In the same way some regularities may have been perceived with the planets, and may have resulted in a certain expectation. Thus in an Assyrian text we read: "Jupiter has stood for a month over its reckoned time" (**11**). Hence its time had been reckoned. The most regular phenomena, such as the periodical invisibility of the planets in the vicinity of the Sun, or, for Jupiter and Saturn, their stately course along the zodiac, making a circumference in 12 and in 29 years respectively, must have become practical experience for the astrologers, just as the sequence of the Moon's figures is practical experience for the savage. We may be sure of it without having express statements in written records. Some centuries later we find texts expressly testifying to knowledge of planetary periods, already at a higher stage. But, in the meantime, conditions had changed considerably.

First politically, when those warring little countries of old had been absorbed into the inner peace of the huge Persian empire, the old omens of lucky for Akkad, unlucky for Elam, lost their sense. Then astronomically, the detection and gradually arising knowledge of the periodicity and the periods must have brought about a new attitude of these observers towards the celestial phenomena. They no longer were the ignorant earthlings anxiously looking at the sky for the messages which the supreme masters of the world should inscribe there. They knew in advance what would be inscribed there, they could predict it, they had some kind of spiritual power over it. Their capacity of prediction gave them a new social qualification, as people who know the ways of the gods. Thus their observing activity impelled them to greater zeal and persistence as a ritual duty in the service of the gods who reveal themselves in these stars. It became more conscious and complete, more detailed and precise; distances of the planets to bright stars were noted numerically, hence perhaps measured, although we know nothing of the instruments used.

The result of this now genuinely astronomical practice appears in the rather scanty inscriptions of the following centuries of rule by the Persian and the Seleucid kings. They show that now the larger periods of the planets were known, the common multiples of synodic period and period of revolution: 8 years for Venus, 71 and 83 for Jupiter, 46 years for Mercury, 47 and 79 years for Mars, 59 years for Saturn; after these intervals the phenomena repeat themselves at nearly the same places. These periods moreover were practically applied for prediction purposes, in such a way that the phenomena, observed the appropriate number of years earlier, were transferred, with small corrections, to the year in question. Thus there exists a tablet entitled: "The first day, the phenomena, the motions and the eclipses for the year 140" (i.e. Seleucid Era) (**12**); it gives phenomena of Jupiter for the years 69 and 57, of Venus for 132, of Mercury for 94, of Saturn for 81, of Mars for 61 and 93: exactly the years in which the data will produce the phenomena for the desired year by addition of the periods named. Ephemerides computed in this way have been found for different years, up to the beginning of our era. Here we have astronomical theory in the form of knowledge of the periods consciously applied in prediction.

A still higher and more perfect form of theoretical knowledge was reached by Babylonian astronomy in the last centuries B.C. Although Babylon had lost its greatness and the Chaldaean country, dominated by Parthian kings, was cut off from the new centres of commerce in the Mediterranean, astronomical activity persisted and now reached its highest stage of perfection before it was extinguished. The tablets dug up from the ruins of old city temples of this time, mostly much-damaged fragments, show nothing but rows of numbers arranged in columns, intermingled with the names of months and zodiacal signs. Their structure and meaning have been unravelled by the work of Father F. X. Kugler, S.J., both an able astronomer and an Assyriologist. In the planetary tables each column contains the successive values of the longitude and the date of one of the special phenomena, opposition positions and heliacal rising and setting. All qualitative description, such as proximity to a star or entrance into a constellation, is absent; there is only pure quantity, given with utmost formal precision in sexagesimal fractions. Now for the first time, because all the data are given by mere numerical values, the inequalities in the planetary motions can become evident as a periodical increase and decrease of the successive intervals. The Chaldaean astronomers did not represent them, as we do, by sinusoidal wave functions based on geometrical construction, but by purely arithmetical zig-zag functions, running with constant differences up and down between an upper and a lower boundary where they are reflected. The same procedure was applied in their lunar tables, where time and place of the conjunctions and oppositions, the first sickle Moon and the eclipses were computed by means of exact data on the Moon's different periods.

We are confronted here not only with far more detailed knowledge of the planetary and lunar motions and their periodic inequalities than in the previous records, but also with its presentation in the form of more abstract scientific theory. The previous form was an excellent device for technical prediction based on the knowledge of periods; but for every prediction reference to former observations was necessary. Here the tables stand by themselves, as pure theory, no longer needing any observational datum. They have not the meaning of mere predictions; the rows of numbers embrace past and future undistinguished and can be extended in both directions indefinitely. It is abstract theory in the form of concrete data; instead of by rules stated in words it is given by the regularities in the tables of figures. The curious fact is that we do not know how it originated; these tables appear suddenly, without intermediate forms that could tell us about their relation to former phases of knowledge. This, however, is only part of the general problem of how the different stages of Babylonian astronomy have developed out of or beside one another and from what observational data, a problem hardly tackled as yet.

Babylonian astronomy in the last centuries presents an admirable system of organized, i.e. theoretical, knowledge. But it is merely formal theory, entirely devoid of any physical interpretation. It is numerical knowledge only; nowhere is the question of structure raised. Its mathematics is only arithmetic, with no geometry. It appears that for these Chaldaean astronomers the planets were not solid bodies moving in orbits through space, but rather, as in olden times, luminaries wandering—although now with unravelled regularity—along the firmament. The cause of this restricted character must be sought for in the fact that this science was entirely in the hands of priests, who by a mighty power of tradition were bound to the old forms of worship and ritual. All their astronomical work

was service to the gods: the most modern planetary tables we find to be introduced by the sacral formula: "In the names of god Bel and goddess Beltis, my masters, an omen." So the way to further theoretical progress, to structural physical theory, was barred. If we consider structural theory an essential characteristic of theoretical science, then for the origin of astronomy one step more must be made. For this we have to turn our attention to another people, to the Greeks.

In living conditions and in character Greece and the Greeks, country and people, in every respect present the greatest possible contrast to Babylonia. In this mountainous peninsula, deeply indented by the sea, with good harbours but with only small bits of arable land separated by wild-wooded ridges, the people remained divided in small communities with their local deities and local priests. They became seafarers and merchants, and many of their expanding numbers emigrated and settled on foreign shores, in Asia Minor and Southern Italy. In such traders and colonists a freer, more independent and more daring spirit developed than among farmers staying at home under permanent conditions; they were less bound to tradition, more open to new ideas. Commerce gave rise to industry, the fabrication of wares for export, and inventiveness became the basis of industrial progress. Here, in the Greek colonies, there arose a ruling class of free and wealthy citizens, the first men in history comparable to the burgher class that in modern times placed its stamp upon European society. Here a new spiritual life revealed itself in a flowering of poetry and of new philosophical world-conceptions.

These first Greek philosophers amaze us as much by the boldness of their ideas about world structure as by the scantiness of their knowledge of astronomical phenomena. From Xenophanes and from Heraclitus—famous even now as the first thinker who saw the world as a process of eternal flux—the opinion is reported that the Sun and the stars were formed anew every day. The cosmical theories of these times were not devised to explain celestial phenomena but their object, as shown by B. Farrington (**13**), was to apply known technical processes to the world at large. These Greek philosophers were thinkers, not observers. Simple phenomena such as the identity of evening and morning star or the course of the planets through the zodiac, which with a little attention could have been noticed by anyone, are attributed as great discoveries to famous philosophers. And this situation remains throughout the centuries in which Greece reached the summit of political power and of literary and artistic brilliancy. Plato, who knows and describes the planets in their sequence, their colour and brightness, says of their wanderings that they are incalculable in multitude and marvellously intricate. Eudoxus' famous theory of the homocentric spheres was an attempt to explain the mere qualitative fact of retrograde motions, which he did by means of a most ingeniously devised system of uniform rotations. But it did not fit the retrogressions as they factually were; obviously there were no observational data to test the details of the theory.

Judging from the amount of their observations and detailed knowledge of the celestial luminaries, the Greeks in their most glorious times cannot be said to have been great in astronomy. But they did something else. They developed geometry and a geometrical mode of thinking as a most admirable system of demonstrable truths, fit to become the basis for a future higher stage of astronomy. They developed a conception of the world, or rather a new way of looking at it, as a system of material bodies and spheres moving and revolving in three-dimensional space. They stood within this world as artificers dominating it

mentally, looking from the outside at the constituting bodies with the Earth itself in the midst, handling them as objects of geometry, cutting their spheres with planes and making the circles and triangles objects of geometrical demonstration. In Aristarchus' determination of the relative distance of Sun and Moon we have a fine example of this attitude towards the heavenly world, Greek astronomy at its best. Yet in the work expounding this method the apparent diameter of Sun and Moon is assumed to be 2°, four times too large.

A new epoch set in when as a consequence of Alexander's conquests the Greek and the Oriental world were merged, the Hellenistic era. Now the Greeks got access to the observing practice and to the abundance of observational knowledge of the Babylonians. The Oriental example gives a new impulse; observations are made at Alexandria and elsewhere. But what we know of it deals only with the fixed stars, not with the planets. And Ptolemy afterwards states that Hipparchus from lack of a sufficient number of observations could not bring his theory of the planets to completion. So far, Greek astronomy is seen to keep its former character, but now the Greek thinkers had at their disposal the numerical results of Babylonian astronomy. Now their geometrical world schemes could be filled in by exact data on periods and inequalities; now the numerical perfection of Chaldaean science could assume the form of geometrical structure in space. The result was the epicycle theory, the supreme product of ancient astronomy.

In the epicycle theory we may be said to have reached our goal, true astronomical science. There is a conception of science which considers only the latest truth as real, true science. It sees the epicycle theory as a primitive erroneous system, to be superseded 1700 years later by the true world system. Looking, however, from an historical point of view we see that it represents correctly the relative motions of the planets in their circular orbits, and that only the original assumptions as to the zero point of the motions had to be corrected afterwards. It is scientific theory in the strictest sense of the word, systematization of observational facts in a world structure suited to the computation and prediction of future events. The study of the history of science presents it to our view as a most important stage in the growth of knowledge, conditioned by the entire social development of the time, as the outcome of observing and thinking activities during a long series of centuries, and as the starting point of later progress towards modern science. Thus we understand how at so early a time, when in no other realm of natural phenomena had knowledge risen above the stage of technical experience, astronomy could rise to the high rank of a science.

References

(1) G. Schiaparelli, *Scientia*, **3**, 1, 1908.
(2) F. K. Ginzel, *Handbuch der math. u. techn. Chronologie*, II, p. 129, 1906.
(3) F. K. Ginzel, *loc. cit.*, pp. 145, 149, 1906.
(4) A. Kötz, *Über die Astronomischen Kenntnisse der Naturvölker Australiens*, 1911.
(5) Hesiod, *Works and Days*, lines 383–384.
(6) A. J. Toynbee, *Civilization on Trial*, Oxford, p. 23, 1948.
(7) G. Smith, *Assyrian Discoveries*, London, 3rd. ed., 1875.
(8) R. C. Thompson, *The Reports of the Magicians and Astrologers of Niniveh and Babylon in the British Museum*, London, 1900.
(9) R. C. Thompson, *loc. cit.*, No. 274 F.
(10) R. C. Thompson, *loc. cit.*, No. 274 A.
(11) R. C. Thompson, *loc. cit.*, No. 183.
(12) F. X. Kugler, S.J., *Sternkunde und Sterndienst in Babel*, I., pp. 59, 85, 1907.
(13) B. Farrington, *Greek Science*, London, 1949; *Proc. Roy. Inst. G.B.*, **33**, 289, 1945.

RADIO EMISSION FROM THE ANDROMEDA NEBULA

R. Hanbury Brown and C. Hazard

(Communicated by A. C. B. Lovell)

(Received 1951 March 29)

Summary

A detailed account is given of the measurement of the radio-frequency radiation from the Great Nebula in Andromeda (M31) using a wave-length of 1·89 m. It is shown that the measured intensity is in good agreement with that to be expected on the assumption that the generation of the radiation is similar to that in the Galaxy. From these results it is estimated that the total extra-galactic radio emission amounts to about 1 per cent of the whole radio flux observed on the Earth.

---

1. *Introduction.*—The experiments of Jansky (1) in 1931 showed that radio-frequency radiation is reaching the Earth from the general direction of the Milky Way. Subsequent measurements by Hey, Parsons and Phillips (2), by Reber (3) and by Bolton and Westfold (4) have shown that the intensity contours of this radiation correspond well with the structure of the Galaxy derived from visual observations. Although the source of the radiation remains unknown, it is generally accepted that the major part of the radiation must be generated by some mechanism which is widespread in the Galaxy.

Attempts have been made in the past to discover whether or not the extra-galactic nebulae generate similar radio emissions. For example, Reber (3) used a paraboloid of 30 feet aperture on a wave-length of 1·87 metres, but failed to find any conclusive evidence of radiation from the nebula M31. During the last few years, however, a much larger paraboloid of aperture 218 feet has been available at the Jodrell Bank Experimental Station. Calculations of the intensity to be expected from M31 based on the assumption that it radiates in a similar way to the Galaxy suggested that with this paraboloid, used in conjunction with the best available receiver, it might be possible to detect the radiation if it existed.

This paper describes the results of an experiment designed to test if such radio-frequency radiation is being emitted by the Great Nebula in Andromeda (M31). A short account of the results has been published previously (5).

2. *Description of apparatus.*—Fig. 1 shows a block diagram of the apparatus. The aerial is a paraboloid of diameter 218 feet and focal length 126 feet. The reflecting surface is formed by long wires which are spaced 8 inches apart and which run parallel to one fixed plane of polarization.

Electromagnetic waves of wave-length large compared with 8 inches and whose polarization is parallel to the wires are reflected by this surface and are focused on to the primary feed which consists of two dipoles and their reflectors mounted on a central mast 126 feet high. This mast is pivoted at its base in such a way that it may be tilted in the north–south plane to 15 degrees on either side of the vertical. The latitude of the system is N. 53° 14′ and it is therefore possible to

direct the beam to declinations between N. 38° and N. 68°. The direction of the beam is estimated by measuring the angle through which the primary feed is displaced from the vertical and applying to this angle a correction established from experiments (**6**) carried out with small paraboloids.

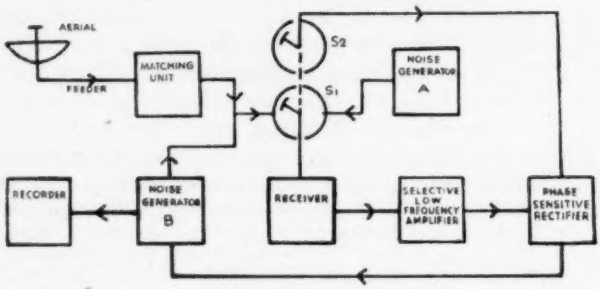


FIG. 1.—*Block diagram of apparatus.*

The gain and beam-width of the aerial have been calculated for various wave-lengths, taking into account the polar diagram of the primary feed, the actual shape of the reflecting surface and the reflection coefficient of the system of wires. They have also been measured at 4·2 metres using an aircraft as a transmitting source. The beam-width at 1·89 metres has been measured by observations of the radio source in Cygnus. The values obtained are shown in Table I.

TABLE I

Characteristics of Aerial System

	72 Mc./s.		158·5 Mc./s.	
	Theoretical	Measured	Theoretical	Measured
Gain (over half-wave dipole)	690	700	2130	...
Beam-width to half power	4° 30′	4° 20′	2° 00′	1° 58′

Fig. 2 shows the shape of the beam measured in both planes by observations of the source in Cygnus, that is with the beam tilted about 14 degrees from the vertical. The broken line in Fig. 2 (*a*) shows the theoretical beam shape when the beam is directed to the zenith.

The aerial system is connected to the receiver through 294 feet of coaxial feeder which has a loss of 2·7 decibels at 1·89 metres.

The receiving equipment is based on that described by Ryle and Vonberg (**7**). Their arrangement has been modified to allow the measurement of aerial temperatures below room temperature. A rotating switch S_1 (Fig. 1) revolves at about 1200 r.p.m. and switches the receiver alternately to a noise generator A and to the aerial in parallel with a noise generator B. Both noise generators are tungsten filament diodes type CV 172. Any difference between the power received in the two positions of the switch appears as a low-frequency modulation of the receiver output. This low-frequency component is amplified by a selective low-frequency amplifier and is applied to a phase-sensitive rectifier operated from a switch S_2 which is synchronous with the aerial switch. The output of the phase sensitive rectifier controls the power delivered from the noise generator B which is in parallel with the aerial. The system automatically maintains a balance such that the power from the aerial plus that from the generator B balances the

power from the generator A. The power from the generator A is kept constant and the variations in the power from the aerial are recorded as variations in power output from the generator B. These variations are recorded continuously by a standard recording milliammeter on a moving chart.

In the region of sky around M31 the effective aerial temperature remained below room temperature and, except when calibrating, the output from the generator A was simply that of a resistance at room temperature. The load impedance of the generator A is accurately pre-set and an adjustable matching unit is used to equalize the impedances of the two branches at the switch S_1. As a further precaution the output of the generator B is calibrated by means of the generator A.

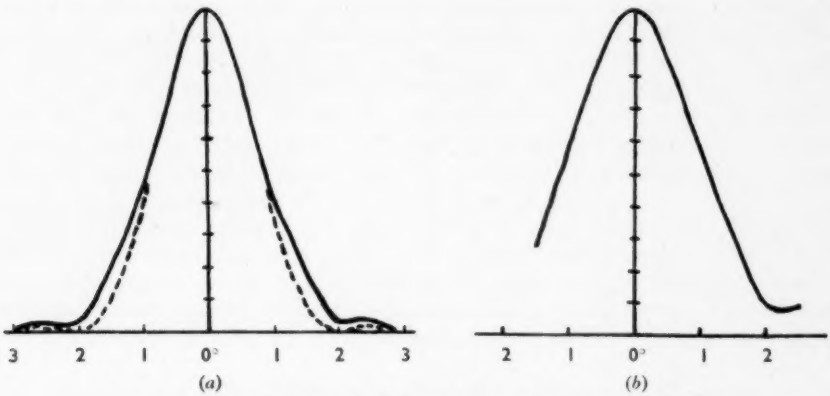


FIG. 2.—*Polar diagrams of 218-ft. aperture paraboloid.* ($\lambda = 1{\cdot}89$ *metres.*)
(*a*) *Beam shape in right ascension.*
(*b*) *Beam shape in declination.*
Ordinates: Intensity in arbitrary units.
Abscissae : Degrees.

The receiver has a midband frequency of 158·5 Mc./s. determined by a crystal-controlled local oscillator. The pre-detector band-width is 1·5 Mc./s., and the radio-frequency amplifier is a "cascode" (**8**) with a noise factor of 3·5. The time constant of response of the whole system is about 10 seconds which is small compared with the 10 minutes taken by a point source at declination N. 40° to transit the aerial beam.

The minimum power incident on the aerial, which can be detected by such an equipment, can be estimated theoretically, assuming that the limit is set by the random fluctuations inherent in the equipment, and not by external interference.

Assuming that

N = noise factor of the receiver,
B = band-width of receiver in cycles per second,
t = time constant of response of the equipment,
T = total effective temperature at input to receiver,
ΔT = r.m.s. fluctuations of output meter,

it can be shown (**9**) that

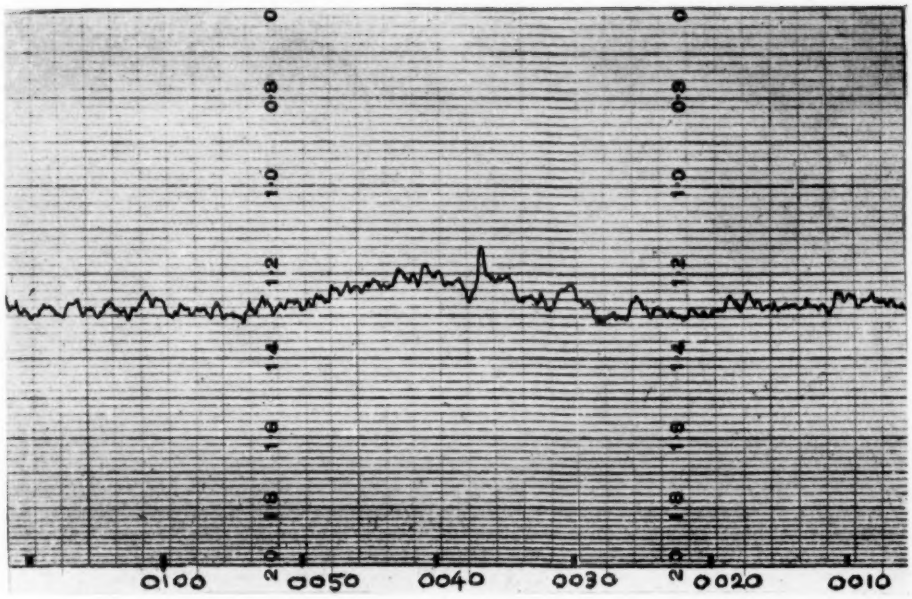
$$\Delta T \sim NT\sqrt{\left(\frac{2\pi}{Bt}\right)}. \qquad (1)$$

Substituting the parameters of the equipment described above, $\Delta T \sim 0{\cdot}7$ deg. C. Practical experience shows that under the conditions of this experiment the minimum deflection which can be detected on a single trace is about twice the r.m.s. fluctuation of the output indicator. Using the theoretical value of ΔT derived above, the minimum change of aerial temperature which can be detected should be about 1·4 deg. C.

Allowing for the loss in the feeder cable and for the decrease in gain of the aerial system when the beam is tilted 14 degrees from the zenith, this value of ΔT corresponds (at 1·89 metres wave-length) to an incident randomly polarized flux of 10^{-25} watts/square metre/c.p.s. Practical experience shows that the minimum detectable flux depends upon the time of day and on the weather. During the day, and in fine weather, the noise level of the equipment is controlled by reception of the "quiet" Sun in minor lobes of the aerial, and by man-made static. Charged rain, electrical disturbances in the atmosphere and radiation from the "disturbed" Sun overload the equipment. During the night the noise level falls off gradually until about midnight and increases again at sunrise. Between midnight and sunrise the minimum detectable flux appears to be controlled sometimes by the theoretical limits discussed in equation (1) and sometimes by unidentified signals which appear to vary with the weather and which probably have a terrestrial origin. On favourable nights it is about 10^{-25} watts/square metre/c.p.s., which is the limit estimated theoretically. By averaging the results of several independent observations it is possible to extend this limit by an amount which has not yet been established by a quantitative experiment. In the present work the limit has been successfully reduced to 5×10^{-26} watts/square metre/c.p.s. by averaging four records.

3. *Method of observation.*—The experiment was carried out by fixing the beam at a number of different elevations corresponding to the region around M31, so that for each diurnal rotation of the Earth the beam swept out a strip of sky 2 degrees wide in declination and 24 hours in right ascension. For each elevation of the beam the intensity of the radio flux received was recorded continuously for about 100 to 200 hours. Altogether, recordings were made for several elevations of the beam, corresponding to declinations between N. 38° 47′ and N. 43° 00′. The variation with right ascension of the flux received at each declination was found by averaging the results obtained on several successive sweeps.

4. *Results.*—The total number of records obtained by the method described above was 90, of which 20 were spoiled by the effects of heavy rain which persisted throughout 1950 August and September. Of the remaining records 50 showed the existence of a localized source of radio-frequency radiation at $00^h\,40^m$ R.A. Plate 4 shows one of these records taken with the beam at declination N. 40° 11′. The remaining 20 were either taken when the beam was directed to declinations N. 38° and N. 43°, or interpretation of the records was made difficult by interference in the period $23^h\,30^m$ R.A. to $01^h\,40^m$ R.A. The curves shown in Fig. 3 were constructed from 30 records which were undisturbed by interference. In order to show clearly the change in intensity with right ascension along each sweep the records have been displaced by an arbitrary amount, corresponding to different declinations, so that the ordinate represents only the relative intensities for the points in each sweep. The sweeps between declinations N. 39° 48′ and N. 41° 58′ indicate a localized source of radiation about $00^h\,40^m$ R.A., superimposed on a slow change of intensity which reaches a minimum between



Facsimile of a single record of the transit of the source in Andromeda taken on 1950 *October* 31 *with the aerial beam directed to declination* 40°11′*N.*
Ordinates : Intensity in arbitrary units.
Abscissae : Hours and minutes of right ascension.

$00^h 20^m$ and $00^h 30^m$ R.A. This slow change corresponds to a gradient of the galactic background radiation, which reaches a minimum when the beam is near its maximum southern galactic latitude.

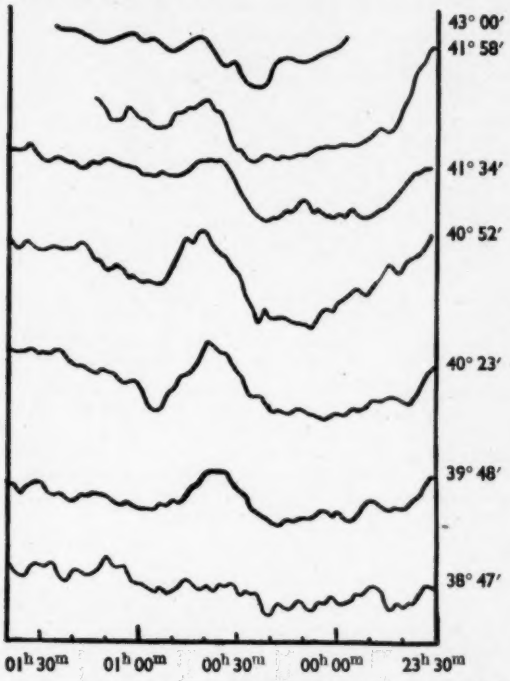


FIG. 3.—*Intensity recorded during transit of the source in Andromeda with the aerial beam directed to different declinations.*

Ordinates: Relative intensity in arbitrary units.
Abscissae: Hours and minutes of right ascension.

These curves cannot be converted into true contours of the intensity from this region because the change of aerial impedance with beam tilt did not permit the difference in absolute intensity between observations at different declinations to be measured reliably. Therefore the background intensity was extrapolated to $00^h 40^m$ R.A. for each curve and the intensity of the background radiation at this point was taken as the zero for each curve when plotting the contours shown in Fig. 4. The contour system (Fig. 4) therefore represents the intensity relative to the background intensity at $00^h 40^m$ R.A., the gradient in declination of the galactic radiation having been removed. The gradient in right ascension has not been removed and so causes a distortion of the contours.

Analysis of the results yields the data shown in Table II. The source has a finite apparent width in right ascension and declination, as can be seen by comparing the apparent widths with the corresponding widths of the intense source in Cygnus, which is at approximately the same declination as the observed source, and whose apparent width is known to be less than 1′·5 of arc (**10**). This comparison eliminates any uncertainty due to possible distortion of the beam. The appropriate curves are shown at (*a*) and (*c*) in Fig. 5. Record (*b*) in Fig. 5

shows the intensity received from another weak source at approximately the same declination as the observed source.

The intensity given in Table II is an integrated value over the source and is twice the intensity observed on one polarization, as it is assumed that the radiation

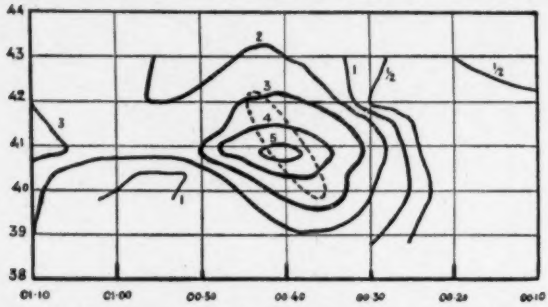


FIG. 4.—*Contours of radio-frequency flux observed near the source in Andromeda with a 2-degree beam.* (1 *unit* $= 10^{-25}$ *watts/square metre/c.p.s.* $\lambda = 1{\cdot}89$ *metres.*)

The contours do not represent the absolute intensity of the radio flux at each point. As explained in the text the gradient of background flux in declination has been removed and the contours show the intensity for each point above the background flux at $00^h\ 40^m$ *R.A. The gradient of background flux in right ascension has not been removed and therefore distorts the contours.*

The broken line shows the outline of the nebula derived from a photograph.

Ordinates: Declination (degrees north).

Abscissae: Right ascension (1 division represents 10 minutes).

TABLE II

Data on Source in Andromeda

Right Ascension	$00^h\ 40^m\ 15^s \pm 30^s$
Declination	40° 50′ N. ± 20′
Apparent widths to half-power	3° (dec.) × 3½° (R.A.)
Apparent widths to half-power of the source in Cygnus	2¼° (dec.) × 2° (R.A.)
Intensity (integrated over source)	10^{-24} watts/sq. metre/c.p.s. ± 25 per cent

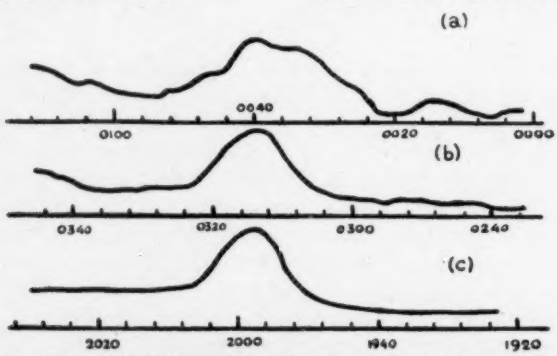


FIG. 5.—*Intensity recorded during transit of three sources through the aerial beam. The scale of intensity for each diagram has been adjusted so that the widths in right ascension can be compared easily.*

(*a*) *Source in Andromeda.*

(*b*) *Weak source in Perseus.*

(*c*) *Intense source in Cygnus.*

Ordinates: Intensity (arbitrary units).

Abscissae: Hours and minutes of right ascension.

is randomly polarized. The previously published (**5**) value of 4×10^{-25} watts/square metre/c.p.s. was calculated assuming the apparent diameter of the source to be small compared with the beam-width, and radiation from the outer regions of the source was neglected. Further measurements and analysis showed the extension of the source to be so great that this assumption was not justified.

The intensity of the source is so small compared with the background radiation and with the random fluctuations in the receiver that the difficulty in interpreting the records is the main cause of errors in the data in Table II. There is no significant systematic error in right ascension as the position of the beam in this coordinate was checked by reference to the intense source in Cygnus whose right ascension is well established. As the declination of the source in Cygnus is not yet known with satisfactory accuracy (**10**) it was not possible to check the position of the beam in declination in the same way. This introduces an additional error into the determination of declination, as the relationship between the angle of tilt of the mast and the angle of the beam is not yet known to better than 1 per cent. Therefore there may be a systematic error of the order of $\pm 10'$ in the measurements of declination. The error in measuring the actual tilt of the mast was negligible.

The error in the measurement of intensity is caused partly by the weakness of the source and partly by inaccuracy in the calibration, these two errors being of comparable magnitude.

5. *Discussion of results.*—The coordinates of the observed source lie close to the centre of the Great Nebula in Andromeda M31 (R.A. $00^h\,40^m$, declination N. $41°\,00'$, Epoch 1950). The apparent dimensions of the source are consistent with a source of radio-frequency radiation of size comparable with the main body of the nebula M31.* It is possible that these results could be due to a fortuitous grouping of two or more point sources with their effective centre coincident with that of M31. From a consideration of the number of observed point sources of intensity comparable to 10^{-24} watts/square metre/c.p.s. and, assuming a random distribution of sources, it can be shown that the probability of observing two sources simultaneously in the beam is about $\frac{1}{4000}$. It is therefore improbable that the finite width of the source is due to a coincidence of two point sources.

The celestial coordinates and finite size of the source indicate with a high degree of probability that the source may be identified with the nebula M31. It is assumed that the two companion nebulae NGC 205 and NGC 221 (which have apparent photographic magnitudes of 10·8 and 9·5 respectively compared with the apparent photographic magnitude 5 of M31 (**11**)) are not contributing substantially to the radiation received.

* *Note added in proof.*—The apparent extension of the source can, in principle, be deduced by applying to the observed results a correction for the finite beam-width of the aerial. However, the low intensity of the source in comparison with the gradient of background radiation precludes an accurate analysis of the extension, particularly in right ascension. The values for the extension of the source given in a previous publication (**5**) were underestimated. A subsequent analysis showed that the observed variation of intensity with declination could be accounted for by a source in which the intensity is a maximum at the centre and decreases to 10 per cent of the central intensity at points separated by 150′. The apparent extension of the source in right ascension appears to be of the same order, but the gradient of background radiation makes this result unreliable.

Inspection of the contour system given in Fig. 4 shows that, when allowance is made for the distortion in right ascension of the contours by the gradient of background radiation, it is possible to interpret the results as due to an elliptical source orientated in a similar manner to M31.

The authors are obliged to Mr J. H. Piddington of the Radiophysics Laboratory, Sydney, whose criticisms have led to the inclusion of this footnote.

6. *Comparison of the radio emission from M31 and the Galaxy.*—From the results described above it appears evident that radio emissions from M31 are being received on the Earth. In this section we compare the measured intensity of this emission with that to be expected if M31 is similar to the Galaxy. An estimate of the intensity to be expected from M31 has been made in two ways, (*a*) by consideration of the luminosities and (*b*) by analysis of the contours of radio flux from the Galaxy.

(*a*) *Comparison of radio flux and light flux.*—This method makes the assumption that the ratio of the radio flux from M31 to the radio flux from a volume near the Sun equals the ratio of the light flux from M31 to the light flux from the same volume near the Sun.

The aggregate luminosity per cubic parsec near the Sun is known to be 0·045 suns. The radio flux from the same volume can be estimated from the temperatures observed (**4**) in the directions of the galactic poles, if it is assumed that the radio flux originates in bodies distributed in a manner similar to the visible stars and that absorption is negligible. The temperatures of the galactic poles have been used because in these directions the estimates of stellar distribution are the most reliable.

Consider an aerial with a beam of solid angle ω, effective area A and gain G directed towards a galactic pole and receiving radiation of wave-length λ metres.

Let

ρ_r = the density of sources per cubic metre at a distance r metres from the Sun,

ρ_0 = the density of sources near the Sun,

$4\pi\alpha$ = the total power radiated by an average source (watts/c.p.s.),

I = the intensity of the radio flux at the Earth (watts/square metre/c.p.s.),

P = the power received by the aerial (watts/c.p.s.),

T = the effective temperature of the aerial (deg. K.),

then
$$I=\int_0^R \frac{\alpha}{r^2}\rho_r \frac{\omega}{4\pi}\cdot 4\pi r^2\,dr=\alpha\omega\int_0^R \rho_r\,dr, \tag{2}$$

where R is the distance from the Sun to the boundary of the Galaxy in a direction perpendicular to the galactic plane, the boundary being taken as the distance at which the stellar density falls to 1 per cent of that near the Sun. Therefore

$$P=kT=\frac{\alpha\omega A}{2}\int_0^R \rho_r\,dr,$$

where k is Boltzmann's constant (the factor $\frac{1}{2}$ is included because the aerial receives radiation in one polarization only), but

$$G=\frac{4\pi A}{\lambda^2}=\frac{4\pi}{\omega}. \tag{3}$$

Hence

$$kT=\frac{\alpha\lambda^2}{2}\int_0^R \rho_r\,dr. \tag{4}$$

The integral $\int_0^R \rho_r\,dr$, evaluated by using the star density/distance tables given by Bok (**12**), equals $1{\cdot}4\times10^{19}\rho_0$.

Taking (**4**) $T=600$ deg. K. and $\lambda=3$ metres,

$$\alpha\rho_0=1{\cdot}3\times10^{-40}\text{ watts/steradian/cu. metre/c.p.s.}$$
$$=3{\cdot}8\times10^{9}\text{ watts/steradian/cu. parsec/c.p.s.,}$$

it has been shown that the law connecting the intensity (I) of the background radiation with wave-length (λ) is $I = \text{constant} \times \lambda^a$, where $0{\cdot}35 < a < 0{\cdot}65$ (**10**).

By assuming $I :: \lambda^{0{\cdot}5}$,

$$\alpha\rho_0 = 3 \times 10^9 \text{ watts/steradian/cu. parsec/c.p.s.}$$

at a wave-length of 1·89 metres. This is the radio flux per steradian from a cubic parsec near the Sun.

The ratio of the light flux from M31 to that from a cubic parsec near the Sun is $2{\cdot}2 \times 10^{10}$. (The luminosity of M31 is taken as 10^9 suns and that of a cubic parsec near the Sun as 0·045 suns.)

Assuming this same ratio for radio-frequency radiation, then the total radio-frequency radiation from M31 is $2{\cdot}2\alpha\rho_0 \times 10^{10}$ watts/steradian/c.p.s. The intensity of this radiation at the Earth should therefore be

$$\frac{2{\cdot}2\alpha\rho_0 \times 10^{10}}{l^2} \text{ watts/square metre/c.p.s.,}$$

where l is the distance of M31 in metres.

If $l = 230{,}000$ parsecs, then $I_{M31} = 1{\cdot}2 \times 10^{-24}$ watts/square metre/c.p.s. which is of the same order as the observed value.

(*b*) *Analysis of contours of radio flux from the Galaxy.*—The contours of radio-frequency radiation over most of the celestial sphere have been measured by Bolton and Westfold (**4**) at a wave-length of 3 metres. These have been used to construct a rough model of the Galaxy as it would appear from the distance of M31.

If absorption is negligible, then the intensity of the radiation observed from this distance will be approximately independent of the orientation of the Galaxy with respect to the observer. In this analysis the Galaxy was assumed to be viewed along the equatorial plane. The observed contours were projected on to a cross-section of the Galaxy through the galactic centre and perpendicular to the observer. The Galaxy was assumed to be a spheroid with dimensions 30,000 × 5,000 parsecs and to have a distribution of mass symmetrical about its minor axis. The projection was made on the assumption that the radiation from any volume of the Galaxy is proportional to the mass in that volume. From these projected contours $\int T(\Omega)\, d\Omega$ was evaluated over the whole cross-section of the Galaxy, where $T(\Omega)$ is the effective temperature observed over an elementary solid angle $d\Omega$ subtended at an observer at a distance of 230,000 parsecs. The intensity (I_G) observed at this distance will be *

$$I_G = \frac{2k}{\lambda^2} \int T(\Omega)\, d\Omega.$$

* This expression can be derived as follows :—

Consider a beam of solid angle ω where ω is large compared with the apparent diameter of the source. The effective temperature (T_e) of this beam will be given by

$$T_e = \frac{\int T(\Omega)\, d\Omega}{\omega}$$

using the same notation as in the text. The power (P) received by the beam is given by

$$P = kT_e = \frac{k\int T(\Omega)\, d\Omega}{\omega} = \frac{I_G}{2} \cdot A,$$

$$I_G = \frac{2k\int T(\Omega)\, d\Omega}{A\omega}.$$

From equation (3) $A\omega = \lambda^2$,

$$\therefore \quad I_G = \frac{2k}{\lambda^2} \int T(\Omega)\, d\Omega.$$

Evaluating the integral as described above, the value of I_G is found to be 9×10^{-24} watts/square metre/c.p.s. at a wave-length of 3 m. Then, assuming $I :: \lambda^{0.5}$ as before, the intensity on the wave-length of 1·89 m. used in the present work should be 7×10^{-24} watts/square metre/c.p.s. for the Galaxy observed at the distance of M31.

By assuming, as above, that the radio flux is proportional to the mass, then the intensity to be expected from M31 can be estimated by comparing its mass with that of the Galaxy. Estimates of the mass of M31 vary from 3×10^{10} to 10^{11} solar masses compared with an estimate of 10^{11} to 2×10^{11} for the Galaxy. The intensity of the radio-frequency radiation from M31 observed on the Earth would thus be expected to be within the range 1×10^{-24} to 7×10^{-24} watts/square metre/c.p.s. which, in fact, includes the observed value.

The agreement between the theoretical values calculated above and the measured intensity suggests that as far as radio-frequency emission is concerned M31 possesses similar characteristics to the Galaxy.

7. *The total radiation from extra-galactic nebulae.*—The contribution to the radio flux incident on the Earth from extra-galactic nebulae has been calculated for a very simple model of the universe, assuming that the ratio of radio flux to light flux found for M31 is true for all nebulae independent of type.

The intensity of the radio flux from M31 observed on the Earth at a wave-length of 1·89 metres is 10^{-24} watts/square metre/c.p.s. and therefore assuming it to be an isotropic radiator the total radio flux emitted is 5×10^{19} watts/steradian/c.p.s. M31 has an absolute visual magnitude of $-17{\cdot}5$ compared with $-15{\cdot}2$ for the average nebula. Therefore the radio flux from an average nebula will be about 5×10^{18} watts/steradian/c.p.s.

If σ = the average space density of nebulae (assumed isotropic),
$4\pi\gamma$ = the total radio flux from an average nebula,
I_E = the intensity at the Earth of extra-galactic radiation per steradian,

then if radiation is received from nebulae in a sphere of radius D and the effects of absorption and recession are neglected,

$$I_E = \gamma\sigma D \qquad \text{(from equation (2))}.$$

Putting $\gamma = 5 \times 10^{18}$ watts/steradian/c.p.s., $\sigma = 1{\cdot}9 \times 10^{-19}$ nebulae per cubic light-year, then for a universe of radius 5×10^8 light-years (corresponding to the range of the 100-inch telescope), $I_E = 4{\cdot}75 \times 10^{-24}$ watts/square metre/steradian/c.p.s. If the radius of the visible universe is taken as $1{\cdot}8 \times 10^9$ light-years, as suggested by the apparent recession of the nebulae,

$$I_E = 1{\cdot}7 \times 10^{-23} \text{ watts/square metre/steradian/c.p.s.}$$

This latter value of I_E corresponds to an effective aerial temperature at 1·89 metres of about 2 deg. K. The average aerial temperature observed over the whole celestial sphere at this wave-length is about 200 deg. K., and this calculation therefore suggests that the extra-galactic radiation contributes about 1 per cent of the total radio flux incident on the Earth.

8. *Acknowledgments.*—The work has been carried out at the Jodrell Bank Experimental Station of the University of Manchester. The aerial system was designed by Dr J. A. Clegg. The construction of the aerial and apparatus was made possible by financial assistance from the Department of Scientific and Industrial Research. We wish to thank Dr A. C. B. Lovell for his interest in

the investigation. One of us (R. H. B.) is indebted to I.C.I. Ltd. for a research fellowship, and the other (C. H.) to the Department of Scientific and Industrial Research for a maintenance grant.

University of Manchester,
Jodrell Bank Experimental Station,
Holmes Chapel, Cheshire:
1951 *March* 28.

References

(1) K. G. Jansky, *Proc. I.R.E.*, **20**, 1920, 1932.
(2) J. S. Hey, J. W. Phillips and S. J. Parsons, *Proc. Roy. Soc.* A, **192**, 425, 1948.
(3) G. Reber, *Ap. J.*, **100**, 279, 1944.
(4) J. G. Bolton and K. C. Westfold, *Aust. J. Sci. Res.* A, **3**, 19, 1950.
(5) R. Hanbury Brown and C. Hazard, *Nature*, **166**, 901, 1950.
(6) S. Silver and C. S. Pao, *Radiation Laboratory Report*, 479, 1944.
(7) M. Ryle and D. D. Vonberg, *Proc. Roy. Soc.* A, **193**, 98, 1948.
(8) H. Wallman, *Proc. I.R.E.*, **36**, 700, 1948.
(9) R. H. Dicke, *Rev. Sci. Instr.*, **17**, 268, 1946.
(10) G. J. Stanley and O. B. Slee, *Aust. J. Sci. Res.* A, **3**, 234, 1950.
(11) H. Shapley and A. Ames, *Harv. Ann.*, **88**, No. 2, 52, 1932.
(12) B. J. Bok, *Distribution of the Stars in Space*, University of Chicago Press, p. 22, 1937.

THE CHEMICAL COMPOSITION OF THE INTERSTELLAR GAS

M. J. Seaton

(Communicated by H. S. W. Massey, F.R.S.)

(Received 1951 June 9)

Summary

Summaries are given of the atomic ionization and recombination data required in the calculation of interstellar ionization equilibrium. The chemical composition of the interstellar gas is discussed from the standpoint of seeking to establish whether there is any evidence for discrepancies between the interstellar and "mean cosmic" abundance ratios. The important *Ca/Na* ratio is the only case for which there is an apparent discrepancy, but it is shown that this may be removed if recent work on the ultra-violet radiation field is accepted as correct.

1. *Introduction.*—In order to obtain information on the chemical composition of the interstellar gas it is necessary to study the ionization equilibrium which is established. Strömgren (1) has shown that the procedure appropriate to the conditions in interstellar space is to equate the rate of photo-ionizations from the ground state of a given ion to the sum of the recombination rates to all excited states. Both of the rate coefficients involved depend upon the continuous atomic absorption cross-sections, and in view of recent work on these it appears to be desirable to tabulate the atomic data in a form which is most convenient for astrophysical applications, and which is independent of any particular assumptions about the conditions prevailing in interstellar space.

We denote by $R_{X'}$ and $R_{X''}$ the number density of ions of a given element X in two consecutive stages of ionization, and by R_e the number density of electrons. The rate of recombinations per unit volume is then given by

$$q^{\text{rec.}} = \alpha R_{X''} R_e. \tag{1}$$

For an electron temperature T_e the radiative recombination coefficient α is given by

$$\alpha = \frac{1}{c^2}\left(\frac{2}{\pi}\right)^{1/2}(mkT_e)^{-3/2}\sum_k \frac{w^k_{X'}}{w^1_{X''}} \exp(I^k/kT_e)\int_{I^k} a^k_\nu (h\nu)^2 \exp(-h\nu/kT_e)d(h\nu), \tag{2}$$

where $w^k_{X'}$, I^k and a^k_ν denote the statistical weight, ionization potential and continuous absorption cross-section of the kth excited state of X' and $w^1_{X''}$ is the ground-state statistical weight of X''.

The rate of photo-ionizations per unit volume is given in terms of the energy density of radiation ρ_ν and the ground-state absorption cross-section a^1_ν by the relation

$$q^{\text{ion.}} = R_{X'}\int a^1_\nu \frac{c}{h\nu}\rho_\nu d\nu = R_{X'}\Gamma. \tag{3}$$

We refer to Γ as the ionization coefficient. Equating the ionization and recombination rates we obtain

$$\frac{R_e R_{X''}}{R_{X'}} = \frac{\Gamma}{\alpha}, \tag{4}$$

which is the Strömgren ionization equation.

2. *Atomic data for the calculation of ionization and recombination rates*

2.1. *The recombination coefficient.*—The recombination coefficients for a number of ions of importance in interstellar space are given in Table I. In calculating the sum of partial recombination coefficients, all states for which no entry has been made in the table have been taken to be hydrogenic. This approximation has been discussed by Bates, Buckingham, Massey and Unwin (**2**); the accuracy of the present atomic data will be discussed in Section 2.3 below. For interpolating the temperature variation of α the most convenient procedure is to plot $\alpha T_e^{1/2}$ against log T_e.

2.2. *The ionization coefficient.*—Estimates of the interstellar radiation field are usually expressed as a sum of black-body energy densities multiplied by corresponding dilution factors,

$$\rho_\nu = \sum_i \delta(T_i)\rho_\nu(T_i). \tag{5}$$

It is therefore convenient to tabulate the atomic data in the form of black-body ionization rates given by

$$\gamma(T_i) = 1{\cdot}00 \times 10^8 \exp(-I/kT_i)\, F(T_i)\,\text{sec}^{-1}, \tag{6}$$

with

$$F(T_i) = 10^{19} \int_0^\infty a_\nu^1 (I+\epsilon)^2 \exp(-\epsilon/kT_i)\, d\epsilon, \tag{7}$$

where a_ν^1 is in cm.^{2}, I and ϵ in Rydbergs (13·5$_8$ e.V.), T_i in deg. K. and k in Rydbergs/deg. K. (i.e. $k = 0{\cdot}638 \times 10^{-5}$). The interstellar ionization rate is then given by

$$\Gamma = \sum_i \delta(T_i)\gamma(T_i)\ \text{sec}^{-1}. \tag{8}$$

The function $F(T_i)$ is given for a number of ions in Table II. In certain cases it is necessary to consider the ionization rate with a cut-off for energies greater than the ionization potential of hydrogen ($(I+\epsilon)>1$); calculations with a cut-off have been marked with an asterisk.

2.3. *Accuracy of the atomic data.*—The accuracy of most of the atomic data should be adequate for treating the problem of interstellar ionization equilibrium (**9**). For the hydrogenic data the correction factor g in Kramers' formula has not been taken fully into account, but the figures given should be correct to within about 10 per cent. For the neutral atoms the only case for which the use of the hydrogenic recombination coefficients is likely to introduce any appreciable error is for *Ca* I, for which the *s*, *p* and *d* series are markedly non-hydrogenic. It is possible that the total α in this case is somewhat underestimated. For the positive ions the departure from hydrogenic values is more marked; in particular it will be seen that the *Ca* II 3*d* recombination term is some six times greater than hydrogenic. Some corrections have been made for *Na* II and *K* II.

In a number of cases it has been necessary to extrapolate the experimental or theoretical cross-section curves. The only case for which this introduces some uncertainty is for the ground state of *K* I. The experimental results of Ditchburn, Tunstead and Yates (**11**) give the cross-section out to 1700 A. ($(I+\epsilon)=0{\cdot}54$ Ryd.), where it is still rising steeply. The uncertainty introduced is greatest for the higher temperature ionization rates, more than 70 per cent of the *K* I ionization coming from the extrapolated part of the curve for T_e greater than 15,000 deg. K. It should be possible to obtain a reliable theoretical cross-section at higher energies using the method of Bates (**16**).

TABLE I

Final ion	Final state	Recombination coefficient α in units 10^{-13} cm.^{3}/sec. $T_e=10^4$	$T_e=10^3$	$T_e=10^2$	Reference
H I	1*s*	0·16	0·50	1·6	(3)
	2*s*	0·020	0·077	0·24	(3)
	2*p*	0·057	0·22	0·69	(3)
	3*s*	0·0059	0·027	0·084	(3)
	3*p*	0·018	0·079	0·25	(3)
	3*d*	0·020	0·092	0·29	(3)
	4*s*	0·0039	0·021	0·065	(4)
	4*p*	0·010	0·053	0·17	(4)
	4*d*	0·012	0·062	0·20	(4)
	4*f*	0·0062	0·033	0·11	(4)
	$\sum_{n=5}^{\infty}$	0·095	0·83	4·9	(1, 2, 5, 6)
	Total	0·41	2·0	8·6	
C I	2*p*	0·16	0·47	1·5	(7)
	Total	0·33	1·7	7·6	
Na I	3*s*	0·0002	0·0013	0·0041	(8, 9)
	3*p*	0·030	0·095	0·30	(10)
	4*s*	0·00004	0·0003	0·0009	(10)
	4*p*	0·009	0·041	0·13	(10)
	Total	0·17	1·2	5·9	
K I	4*s*	0·001	$0{\cdot}000_3$	0·001	(11)
	4*p*	0·025	0·039	0·12	(12)
	Total	0·16	1·1	5·6	
Ca I	$4s^2$	0·031	0·10	0·33	(13)
	Total	0·17	1·2	6·0	
He II	1*s*	0·65	2·1	6·6	(3)
	2*s*	0·087	0·31	0·98	(3)
	2*p*	0·26	0·86	2·8	(3)
	3*s*	0·033	0·11	0·33	(3)
	3*p*	0·10	0·32	1·0	(3)
	3*d*	0·12	0·37	1·2	(3)
	4*s*	0·021	0·083	0·26	(4)
	4*p*	0·056	0·21	0·68	(4)
	4*d*	0·064	0·25	0·78	(4)
	4*f*	0·034	0·13	0·42	(4)
	$\sum_{n=5}^{\infty}$	0·74	5·0	25	(1, 2, 5, 6)
	Total	2·2	9·7	40	
Na II	2*p*	0·19	0·59	1·9	(14)
	$n=3$	0·50	1·6	5·0	See text
	Total	1·6	7·9	34	
K II	3*p*	0·39	1·2	3·8	(5)
	3*d*	0·69	2·2	6·9	See text
	Total	2·0	9·1	38	
Ca II	3*d*	0·69	2·2	6·9	(9, 15)
	4*s*	0·003	0·007	0·024	(9, 13, 15)
	4*p*	0·17	0·55	1·7	(15)
	Total	1·7	8·1	35	

TABLE II

Atomic data for the calculation of ionization rates

Ion	I in Ryd.	$F(T_i)$ F(50,000)	F(25,000)	F(15,000)	F(10,000)	F(5,000)
H I	1·000	16·0	9·0	5·6	3·8	2·0
C I*	0·828	12·0	9·4	7·0	5·1	2·6
Na I	0·378	0·0081	0·0067	0·0056	0·0046	0·0031
K I*	0·319	0·42	0·11	0·026	0·0071	0·00091
Ca I	0·449	10·4	6·1	4·0	2·7	1·6
Na II	3·47	330·0	150·0	86·0	56·0	28·0
K II	2·34	640·0	300·0	170·0	120·0	57·0
Ca II	0·873	0·57	0·21	0·11	0·067	0·031
Ca II*	0·873	0·13	0·095	0·073	0·055	0·030

For references to atomic data see Table I.
The asterisk denotes cut-off for $(I+\epsilon)>1$.

3. *Interstellar ionization equilibrium.*—To make further progress we must make a definite choice for the interstellar radiation field. In Table III we give the ionization rate Γ and the equilibrium constant Γ/α calculated using the temperatures and dilution factors given by Dunham (**17**). In using these data there are four possible sources of error to be borne in mind:

(i) Dunham states that his results are of a preliminary nature and that further work may call for some revision.

(ii) The correction for general galactic absorption and scattering is rather uncertain; we have used Dunham's data without an absorption correction.

(iii) The ultra-violet radiation field is obtained using assumed black-body curves for stars of representative spectral classes. This source of error will be discussed in Section 4.3 below.

(iv) In discussing the conditions in a single cloud, allowance must be made for the possibility of marked local variations from a statistical field due to the influence of nearby stars.

TABLE III

Ionization equilibrium data using the radiation field of Dunham†

Ion	Γ in 10^{-12} sec.$^{-1}$	Γ/α in cm.$^{-3}$ $T_e=10^4$	$T_e=10^3$	$T_e=10^2$
H I	43	105	22	5·0
C I*	110	320	63	14
Na I	4·0	24	3·3	0·68
K I*	42	260	38	7·5
Ca I	1400	8200	1200	230
Na II	0·62	0·39	0·079	0·018
K II	27	13	3·0	0·71
Ca II	2·5	1·5	0·31	0·071
Ca II*	0·86	0·51	0·11	0·025

As before, an asterisk denotes a cut-off for $(I+\epsilon)>1$.

† Some of the data in this table may be compared with a similar table given by Strömgren (**1**). The only marked discrepancy is for *Ca* II, for which Strömgren gives values of Γ/α about four times greater than ours. This is probably due to the use of the cross-section of Bates and Massey (**13**) without correction for a normalization error pointed out by Green (**15**).

The opportunity is taken of correcting two errors in a previous paper (**5**) by the present author. Firstly, in Table III, $T_e=10^2$, 10^3, 10^4 should read $T_e=10^4$, 10^3, 10^2. Secondly, an incorrect multiplying constant has been used in calculating the *K* II recombination coefficient. Corrected values are given in the present paper.

4. *The chemical composition of the interstellar gas*

4.1. *Introduction.*—There are not at present sufficient data available to make possible an unambiguous determination of the chemical composition of the interstellar gas. It would seem reasonable to expect, however, that the composition would be the same, to a first approximation, as that of the Sun, stars, meteors and gaseous nebulae, as expressed in tables of "mean cosmic abundances". We will therefore adopt the approach of seeking to establish whether there is any evidence for systematic departures from the mean cosmic abundance ratios. In order to have a definite basis for discussion we adopt as standard the cosmic abundances given in a recent review by Harrison Brown* (**18**); values for the elements of interest are quoted in Table IV.

TABLE IV

Mean cosmic abundance ratios (after Harrison Brown)

	Relative number of atoms		
H	10^6	*K*	0·20
C	230	*Ca*	1·9
Na	1·3	*Fe*	52
Mg	25	*Ti*	0·074
Si	29		

$$C+Mg+Si+Fe=336$$

Assuming the composition of the interstellar gas to be the same as the mean cosmic ratios suggest, the electron density in regions of ionized hydrogen (*H* II regions) may be equated to the hydrogen density, and in regions of neutral hydrogen (*H* I regions) it may be equated to the combined densities of *C*, *Mg*, *Si* and *Fe*. A guide to the temperature is provided by the recent work of Spitzer (**20**), which suggests that the electron temperature is of order 10^4 deg. K. in *H* II regions and 10^2 deg. K. in *H* I regions. For the ionization rates in *H* I regions we take, where necessary, a cut-off for $(I+\epsilon)>1$.

4.2. *The sodium density.*—We consider first the observational results for the density of neutral sodium atoms. For an *H* I region most of the sodium will be singly ionized, and the total sodium density is therefore given by

$$R_{Na}=R_{Na\,II}=R_{Na\,I}\cdot\frac{1}{R_e}\cdot\left(\frac{\Gamma}{\alpha}\right)_{Na\,I}.$$

Assuming cosmic abundance ratios, the hydrogen concentration is then given by

$$R_H^2=\left(\frac{R_H}{R_{Na}}\right)\left(\frac{R_H}{R_e}\right)R_{Na\,I}\left(\frac{\Gamma}{\alpha}\right)_{Na\,I}$$

$$=1{\cdot}6\times10^9 R_{Na\,I}\quad(T_e=100\text{ deg. K.}).$$

From a very extensive analysis of the sodium D absorption lines Spitzer (**21**) deduces the concentration of neutral sodium atoms per cm.^{3} to be 5×10^{-8} in the more opaque clouds and 2×10^{-8} in the less opaque clouds.† These figures give a hydrogen concentration in the clouds of 9/cm.^{3} and 6/cm.^{3}, which is in good agreement with the value derived by Strömgren (**1**) from the extent

* The problem of the chemical composition of the gaseous nebulae has been further discussed by Aller (**19**).

† It is to be noted that the neutral sodium would be expected to be strongly concentrated in clouds (proportional to the square of the hydrogen concentration). We therefore consider Spitzer's data for the clouds rather than that for an average over all space.

of hydrogen emission regions. Our discussion of this point follows closely the treatment of Strömgren.

It is to be noted that the derived hydrogen density is proportional to the square root of the ionization function, and will therefore be insensitive to the approximation used for the radiation field. The resulting electron concentration is of order 0·003/cm.^{3}; carbon, which gives most of the free electrons, is readily verified to be almost completely ionized under these conditions. The hypothesis that the absorbing clouds are *H* II regions leads to a hydrogen concentration of order one atom per cm.^{3} in the clouds, which appears to be too low.

4.3. *The Ca/Na ratio.*—Many writers have commented on an apparent discrepancy between the interstellar value for the *Ca/Na* ratio and the mean cosmic and terrestrial values.* It is of particular interest to see whether this discrepancy is real. It appears to be well established that the interstellar *Ca* II/*Na* I ratio is of order unity. Strömgren (**1**) obtained the value 0·6 as a mean for seven stars. In Table V we give results for a further 14 stars suitable for analysis by the doublet ratio method (**1**), and common to lists of equivalent widths for *Na* D lines given by Spitzer (**21**), and of equivalent widths for *Ca* II H and K given by Spitzer, Epstein and Li Hen (**24**).

TABLE V

Star HD number	$\frac{R_{Ca\,II}}{R_{Na\,I}}$	Weight	Star HD number	$\frac{R_{Ca\,II}}{R_{Na\,I}}$	Weight
1337	1·3	0·9	184915	0·71	2·0
35619	0·36	2·0	185780	2·2	1·0
163472	0·29	0·5	186994	0·39	1·0
164353	0·31	1·0	188209	0·29	0·3
167264	0·83	1·0	189550	0·40	0·7
167838	0·56	0·5	190919	2·2	0·1
175876	0·71	1·0	204172	1·3	3·0

Weighting factors have been assigned according to the weights given in the equivalent width tables, with an additional factor depending on the part of the curve of growth involved. The weighted mean gives $R_{Ca\,II}/R_{Na\,I}=0{\cdot}85$, in good agreement with Strömgren's value. We adopt the value 0·8 for future calculations. There appears to be little doubt as to its essential correctness.

For an *H* I region with $T_e=100$ deg. K. the ionization data in Table IV give

$$\frac{R_{Ca}}{R_{Na}}=0{\cdot}8\left\{\frac{R_e+0{\cdot}025}{R_e+0{\cdot}68}\right\}.$$

For R_e of the order of 0·003/cm.^{3} this gives $R_{Ca}/R_{Na}=0{\cdot}033$, which is much less than the mean cosmic value of 1·5. The result is quite insensitive to the exact value of the electron concentration and to the assumed electron temperature.†

There is a much greater sensitivity to the radiation field, since the *Na* I and *Ca* II ionization rates depend on quite different spectral regions (2413 – 1800 A. for *Na* I and 1044–911 A. for *Ca* II). This may be expressed in terms of the temperatures of the stars responsible for the ionizations. For *Na* I 80 per cent of the ionization comes from stars with $T_i \leqslant 16{,}000$ deg. K., while for *Ca* II only 20 per cent of the ionization comes from these stars.

* Eddington (**22**), Struve (**23**), Dunham (**17**), Strömgren (**1**).

† For an *H* II region we obtain $R_{Ca}/R_{Na}=0{\cdot}06$ for $R_H=1$. For $R_H=10$/cm.^{3} the ratio is 0·25, but this gives the *Na/H* ratio to be too great by a factor of 100.

The possible inaccuracies in the atomic data are certainly too small to account for the discrepancy (**9**), and it is therefore necessary either to accept an interstellar *Ca*/*Na* ratio much less than the mean cosmic value, or to make a rather drastic alteration in the assumed radiation field. There is some evidence that the second alternative might be justified. Firstly, for the Sun it is possible to calculate the *Na* I ionization rate using ultra-violet intensities obtained from rocket spectrograms, and it is found that the result is about one-tenth of that obtained from the conventional 6000 deg. K. black-body model (**25**). Secondly, for stars of spectral type around B0, which give the major contribution to the *Ca* II ionization, it is possible to calculate the ultra-violet flux using the method of model atmospheres. The results of Underhill (**26**) and of Pecker (**27**) give an ionization rate which is some seven times the value obtained using the effective temperatures adopted by Dunham. If these changes can be regarded as representative, the previously calculated interstellar *Ca*/*Na* ratio must be multiplied by a factor of order 70, giving $R_{Ca}/R_{Na}=2{\cdot}3$, which removes any discrepancy between the mean cosmic and interstellar values.

4.4. *The interstellar lines in* χ^2 *Orionis.*—The spectrum of χ^2 Orionis is particularly rich in interstellar absorption lines. In the first column of Table VI we give the ionic abundances as derived by Strömgren (**1**) from data of Dunham (**17**). These are calculated on the assumption that the absorbing path-length, D, is 15 parsecs, which is of the order of the diameter of a single cloud.*

TABLE VI

Abundances from χ^2 *Orionis absorption lines*

Ion densities $D=15$ psc.	Atom	*H* I region, $T_e=100$ deg. K.				*H* II region, $T_e=10^4$ deg. K.	
		$D=15$ psc.		$D=25$ psc.		$D=15$ psc.	
		χ^2 Orionis	Cosmic	χ^2 Orionis	Cosmic	χ^2 Orionis	Cosmic
	H	...	70	...	40	...	6
Na I 2×10^{-6}	*Na*	60×10^{-6}	90×10^{-6}	60×10^{-6}	50×10^{-6}	10×10^{-6}	8×10^{-6}
K I 3×10^{-9}	*K*	10×10^{-6}	14×10^{-6}	10×10^{-6}	8×10^{-6}	4×10^{-6}	1×10^{-6}
Ca I 2×10^{-9}	*Ca*	40×10^{-6}	130×10^{-6}	60×10^{-6}	80×10^{-6}	4×10^{-6}	10×10^{-6}
Ti II 5×10^{-6}	*Ti*	5×10^{-6}	5×10^{-6}	3×10^{-6}	3×10^{-6}	$>5\times10^{-6}$	$0{\cdot}4\times10^{-6}$

For an *H* I region we may assume that nearly all the titanium is singly ionized (the ionization energies are 0·502 Ryd. for *Ti* I and 1·005 Ryd. for *Ti* II). Equating the total *Ti* density to the observed *Ti* II density, and assuming cosmic abundances, gives $R_H=70/\text{cm.}^3$ and $R_e=0{\cdot}023/\text{cm.}^3$. That these values are greater than would be expected for an average cloud is not surprising since the χ^2 Orionis lines are abnormally strong. Using an electron density of $0{\cdot}023/\text{cm.}^3$ the concentrations of *Na*, *K* and *Ca* may be calculated. The results given in Table VI are seen to be in reasonable agreement with the mean cosmic values (fitted to agree with *Ti*). For an assumed path-length of 25 parsecs the agreement is seen to be almost perfect; the same agreement would be obtained using a path-length of 15 parsecs and increasing all ionization rates by a factor

* The high-dispersion measurements of Adams (**28**) suggest that the absorption is almost entirely due to a single cloud.

of 25/15. If the radiation field were changed in the manner suggested in Section 4.3, it would still be possible to obtain the *Ti*, *K* and *Na* abundances in agreement with the cosmic values, provided a much greater path-length was assumed, but the *Ca*/*Na* ratio would be a good deal greater than the cosmic value. It is not necessary, of course, to assume that the field for a single cloud is the same as the statistical field for many clouds.

We have also considered in Table VI the hypothesis that the absorbing cloud is an *H* II region. For this case it is only possible to give a lower limit for the *Ti* density, since there would be an appreciable amount of *Ti* III. We have taken $R_e = R_H = 6/\text{cm.}^3$, which gives the best fit for *Na*, *K* and *Ca*. It will be seen that the agreement with the cosmic values is not so satisfactory as on the *H* I hypothesis.*

4.5. *The Ca* II/*Ca* I *ratio.*—Observations of the *Ca* II/*Ca* I ratio would be valuable in leading to estimates of the electron concentration which would not involve any assumptions about the absorbing path-length or the chemical composition of the gas. Unfortunately, when the *Ca* I lines are observable the *Ca* II lines are nearly saturated and an accurate determination of the number of atoms is very difficult. Struve (**23**) has estimated the ratio to be between 100 and 150 for 55 Cygni and χ^2 Orionis†, using *assumed* equivalent widths and curves of growth. What appears to be a more reliable estimate for χ^2 Orionis has been made by Dunham (**17**), who gives the ratio to be 3500. For an *H* I region ($T_e = 100$ deg. K.) this gives an electron concentration of $0{\cdot}066/\text{cm.}^3$, which in view of the uncertainties involved, does not compare unfavourably with our previous estimate of $0{\cdot}023/\text{cm.}^3$ ($D = 15$ psc.). For an *H* II region the corresponding value is $2{\cdot}3/\text{cm.}^3$.

5. *Discussion and conclusions.*—It appears to be well established that the relative abundances of the elements in the interstellar gas are of at least the same approximate order of magnitude as for other cosmic sources. The agreement is rather better on the assumption that the average cloud is an *H* I region than that it is an *H* II region. The only case for which there is some evidence of a departure from cosmic values is for the *Ca*/*Na* ratio, but in this case it appears at least plausible that a revised estimate of the ultra-violet radiation field might clear up the discrepancy. Our results apply most directly to the metals, *Na*, *K*, *Ca* and *Ti*, and indirectly to the elements, mostly *C*, which supply the free electrons. Hydrogen densities derived from other considerations are in good agreement with values derived on the assumption that the composition is in agreement with mean cosmic values.

In discussing the chemical composition of interstellar *matter* it is necessary to take into account the composition of the grains as well as that of the gas. According to van de Hulst (**29**), the grains constitute 0·6 per cent of the total mass of interstellar matter, while for cosmic abundances the metals contribute 0·3 per cent of the total mass. Appreciable departures from cosmic abundances in the gas might therefore be expected if there were a marked concentration of metals in the grains. However, recent work‡ suggests that dielectric grains,

* Strömgren (**1**) has considered the possibility of the *Na*, *K* and *Ca* lines originating in an *H* II region, and the *Ti* lines originating in an *H* I region. However, it would appear that an *H* I region of sufficient extent to account for the *Ti* II lines would also be sufficient to account for all the other lines.

† Bates and Massey (**13**) have considered the theoretical implications of this.

‡ H. C. van de Hulst (**29**), Greenstein (**30**), Whitford (**31**).

composed principally of hydrogen and oxygen, are more probable than metallic grains, and this is consistent with the suggestion that the abundance of metals in the gas is in substantial agreement with mean cosmic values.

Department of Physics,
University College, London:
1951 *June* 8.

References

(1) B. Strömgren, *Ap. J.*, **108**, 242, 1948.
(2) D. R. Bates, R. A. Buckingham, H. S. W. Massey and J. J. Unwin, *Proc. Roy. Soc.* A, **170**, 322, 1939.
(3) D. R. Bates, *M.N.*, **106**, 432, 1946.
(4) Landolt-Börnstein, *Physical Tables*, Band 1, Teil 1, p. 260, Aufl. 6, Springer-Verlag, Berlin, 1950.
(5) M. J. Seaton, *M.N.*, **110**, 247, 1950.
(6) L. Spitzer, *Ap. J.*, **107**, 6, 1948.
(7) D. R. Bates and M. J. Seaton, *M.N.*, **109**, 698, 1949.
(8) R. W. Ditchburn and P. J. Jutsum, *Nature, Lond.*, **165**, 723, 1950.
(9) M. J. Seaton, *Proc. Roy. Soc.* A, **208**, 408 and 418, 1951.
(10) M. Rudkjøbing, *Publ. Kbh. Obs.*, **18**, 1, 1940.
(11) R. W. Ditchburn, J. Tunstead and J. G. Yates, *Proc. Roy. Soc.* A, **181**, 386, 1943.
(12) B. Mishra, *Thesis*, London, 1948.
(13) D. R. Bates and H. S. W. Massey, *Proc. Roy. Soc.* A, **177**, 329, 1941.
(14) D. R. Bates, *M.N.*, **106**, 423, 1946.
(15) L. C. Green, *Ap. J.*, **109**, 289, 1949.
(16) D. R. Bates, *Proc. Roy. Soc.* A, **188**, 350, 1946.
(17) T. Dunham, *Proc. Am. Phil. Soc.*, **81**, 277, 1939.
(18) Harrison Brown, *Rev. Mod. Phys.*, **21**, 625, 1949.
(19) L. H. Aller, *Ap. J.*, **111**, 609, 1950.
(20) L. Spitzer, *Ap. J.*, **111**, 593, 1950.
(21) L. Spitzer, *Ap. J.*, **108**, 274, 1948.
(22) A. S. Eddington, *Proc. Roy. Soc.* A, **111**, 424, 1926.
(23) O. Struve, *Proc. Nat. Ac. Sci.*, **25**, 36, 1939.
(24) L. Spitzer, I. Epstein and Li Hen, *Ann. d'Astrophys.*, **13**, 147, 1950.
(25) D. R. Bates and M. J. Seaton, *Proc. Phys. Soc.* B, **63**, 129, 1950.
(26) A. B. Underhill, *Ap. J.*, **111**, 203, 1950.
(27) J.-C. Pecker, *Ann. d'Astrophys.*, **13**, 433, 1950.
(28) W. S. Adams, *Ap. J.*, **109**, 354, 1949.
(29) H. C. van de Hulst, *Rech. Astr. de l'Obs. de Utrecht*, XI, part 2, 1949.
(30) J. L. Greenstein, *Ap. J.*, **104**, 403, 1946.
(31) A. E. Whitford, *Ap. J.*, **107**, 102, 1948.

APPROXIMATE INTEGRATION OF THE EQUATION OF TRANSFER

J. B. Sykes *

(Communicated by the Director, University Observatory, Oxford)

(Received 1951 April 3)

Summary

The value of numerical integration in obtaining approximate solutions of an equation of transfer, and the different methods at our disposal, are discussed. It is shown that although the Newton-Cotes method, used by Kourganoff, is better than the Gauss method, used by Chandrasekhar, both are inferior to a new method, the double-Gauss, discovered by the author. The errors in the approximate values of the source-function and the limb-darkening in all three methods are tabulated for various approximations, and illustrated by graphs.

Numerical integration is a device to which we must turn when a function has to be integrated but its indefinite integral is unknown. This may arise either because the function cannot be integrated in elementary terms, or because it is itself unknown. Numerical integration consists in replacing the integral by the sum of the values of the integrand at a number of discrete points of the interval, each suitably weighted.

If the function is known, or its graph is available, these values are easily found. In radiative transfer problems, however, nothing is known of the function $I(\tau, \mu)$—the intensity of radiation at optical depth τ and in a direction making an angle arc $\cos\mu$ with the outward normal—except that it satisfies the equation of transfer, for instance

$$\mu \frac{dI(\tau,\mu)}{d\tau} = I(\tau,\mu) - \tfrac{1}{2}\int_{-1}^{1} I(\tau,\mu')\,d\mu'$$

for a plane-parallel grey atmosphere in radiative equilibrium and local thermodynamic equilibrium (the "standard case") or a plane-parallel isotropically-scattering grey atmosphere.

However, if the integral on the right can be replaced by the sum

$$\sum_{j=1}^{n} a_j I(\tau,\mu_j),$$

where the μ_j are a finite number of points in the interval $(-1, 1)$, then we obtain, by replacing μ by each μ_j in turn, a set of linear differential equations with constant coefficients, which are easily solved to give $I(\tau, \mu_j)$. This procedure, which was first suggested by Wick (**9**), amounts to dividing the radiation field into distinct beams in the directions μ_j, as was done for two beams by Schwarzschild (**8**) in his original approximate solution of the equation of transfer. The solutions are of the form†

$$I(\tau,\mu_i) = \tfrac{3}{4}F\left[\sum_{\alpha=1}^{\nu-1} \frac{L_\alpha \exp(-k_\alpha\tau)}{1+\mu_i k_\alpha} + \mu_i + Q + \tau\right],$$

* Skynner Senior Student in Astronomy, Balliol College, Oxford.

† In this and the next two equations, n is even and $\mu_1, \ldots, \mu_n$ have been renumbered $\mu_{-\nu}, \ldots, \mu_\nu$, with $\mu_{-\tau} = -\mu_\tau$.

where k_α is a positive root of the equation

$$\sum_{j=-\nu}^{\nu} \frac{a_j}{1+\mu_j k_\alpha} = 2,$$

and the L_α and Q satisfy the boundary conditions

$$\sum_{\alpha=1}^{\nu-1} \frac{L_\alpha}{1-\mu_i k_\alpha} + Q = \mu_i \ (i = 1, \ldots, \nu),$$

expressing that there is no radiation incident on the surface of the atmosphere. These results were first obtained by Chandrasekhar (**1a**).

We must now consider how the points of division and the weights are to be chosen. If $F(\mu)$ is the function

$$\prod_{j=1}^{n} (\mu - \mu_j),$$

where μ_j are the points of division, then

$$\phi(\mu) \equiv \sum_{j=1}^{n} I(\mu_j) \frac{F(\mu)}{(\mu-\mu_j)F'(\mu_j)}$$

is Lagrange's polynomial of degree less than or equal to $n-1$ taking the same values as $I(\mu)$ for $\mu = \mu_j$, and

$$\int_{-1}^{1} \phi(\mu)\, d\mu = \sum_{j=1}^{n} a_j I(\mu_j),$$

where

$$a_j = \frac{1}{F'(\mu_j)} \int_{-1}^{1} \frac{F(\mu)}{\mu - \mu_j}\, d\mu.$$

Thus if the μ_j are given, the weights a_j can be calculated once and for all. There are now three chief methods of dividing the interval $(-1, 1)$:—

1. *Newton-Cotes method.*—In this, the oldest system (**6**), the interval is divided into equal parts, there being "open" and "closed" divisions according as the end-points ± 1 are omitted or included. The weights are

$$a_j = \int_{-1}^{1} \frac{\prod_{r \neq j} (\mu - \mu_r)}{\prod_{r \neq j} (\mu_j - \mu_r)}\, d\mu.$$

This method stands alone when the graph of the function is given, so that the values corresponding to a given μ can be measured; for example, the indicator diagram of a steam engine.

2. *Chebyshef method.*—Here the weights are taken equal, and the division is given by the zeros of the polynomial part of

$$x^n \exp\left[-\frac{n}{2.3x^2} - \frac{n}{4.5x^4} - \frac{n}{6.7x^6} - \ldots \right].$$

This method is useful in the case of values equally liable to error, for the best result is obtained by equal weighting, but has the disadvantage that for high values of n some weights become imaginary.

3. *Gauss method.*—Both the above methods integrate exactly with n points of division any polynomial of degree less than n. Gauss (**2**) in 1814 inquired how the μ_j should be chosen so as to integrate exactly polynomials of the highest degree possible. Since there are n roots and n weights at our disposal, this

degree is $2n-1$. Because $f-\phi$ is zero for $\mu=\mu_j$, where ϕ is defined in terms of $f(\mu_j)$, we have

$$f(\mu)-\phi(\mu)=F(\mu)\sum_{l=0}^{n-1} q_l\mu^l$$

if f is of degree $2n-1$, and the integration is exact if and only if

$$\int_{-1}^{1} F(\mu)\mu^l\,d\mu=0 \text{ for } l=0,\ 1,\ \ldots,\ n-1.$$

It can be shown that this is uniquely satisfied by $F(\mu)\equiv P_n(\mu)$, the Legendre polynomial of degree n. The division points are therefore the n zeros of $P_n(\mu)$ in $(-1,\ 1)$, and the weights are given by

$$a_j=\frac{1}{P_n'(\mu_j)}\int_{-1}^{1}\frac{P_n(\mu)}{\mu-\mu_j}\,d\mu.$$

This last method was the one selected by Chandrasekhar (**1b**) in dealing with the transfer problem mentioned above, and in calculating numerical values from his results. A variant of it, due to Radau, has been used by Kopal (**3**) in the reduction of eclipse observations of the Sun's limb; here $F(\mu)\equiv(1-\mu^2)P_n'(\mu)$, and has zeros at $\mu=\pm 1$, which makes it more suitable for his purposes.

Kourganoff (**4b**) has recently challenged the apparently obvious superiority of Gauss's formula, on the following grounds:—

(1) $I(\tau,\ \mu)$ is not representable by polynomials.

(2) The zeros of $P_n(\mu)$ are crowded towards $\mu=1$, where I is varying slowly, and spaced out towards $\mu=0$, where I varies rapidly.

In proof of (1), Kourganoff insists that $I(0,\mu)$ is discontinuous at $\mu=0$ and so cannot be represented by a polynomial. This, however, does not appear to hold as soon as we penetrate beneath the surface, as the many approximations to $I(\tau,\mu)$ by series in μ^r or in $P_r(\mu)$ have shown.

(2) is true enough in the case of even-order Legendre functions, which have been exclusively used by Chandrasekhar. The odd-order functions have a zero at $\mu=0$, but this only leads to further difficulties. The function $q(\tau)$, defined by

$$J(\tau)\equiv\tfrac{1}{2}\int_{-1}^{1} I(\tau,\mu)\,d\mu=\tfrac{3}{4}F[\tau+q(\tau)],$$

is known to take the value $1/\sqrt{3}$ for $\tau=0$—the Hopf-Bronstein relation. If, however, the division at $\mu=0$ is given the weight a_0, it can be shown that q_0 is given as

$$\frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{(1-a_0/2)}},$$

and this cripples the approximation from the start.

Kourganoff proposes instead to apply to the radiative transfer problem the Newton-Cotes method, with points of division evenly spaced, which at any rate meets objection (2). He calculates the function $q(\tau)$ for various values of τ in the second, third and fourth approximations (namely those with four, six and eight points in the whole interval) and also the error

$$\frac{(\tau+q)_{\text{calc}}-(\tau+q)_{\text{exact}}}{(\tau+q)_{\text{exact}}},$$

and compares them with the Gaussian values in similar approximations. (The exact solution for q, by which we judge the error, has been found by Mark (**5**) by means of symbolic calculus, and by Kourganoff (**4a**) by variational methods.) It turns out that the Newton method is better in the even approximations near the surface ($0<\tau<1$), the essential region according to Kourganoff, but is slightly worse in the third. The success of Newton's method is in part due to the changing sign of the error, which tends to keep it near zero; on the other hand, this brings about a serious difficulty, in that the error in the fourth approximation is greater than that in the second over $\tau=0{\cdot}4$ to $0{\cdot}6$, a rather wide range, while the error in the third over this range is greater than either! In Gauss's method, on the contrary, every error curve is simply a diminished version of the last, and one can at least be sure that by going to a higher approximation one is getting a more accurate result. The expression of the error in this form, though justifiable, also tends to conceal the relative inaccuracy of $q_\infty(=Q)$ in the low Newton approximations. A further objection to Newton's method is the following:—

In the course of the solution of the differential equations, it is necessary to express the flux integral

$$\int_{-1}^{1} I(\tau,\mu)\mu\, d\mu$$

also as a sum over discrete values μ_j. Chandrasekhar puts this equal to

$$\sum_{j=1}^{n} a_j\mu_j I_j,$$

where the a_j are the same weights as before, whereas according to the strict theory of numerical integration it should be equated to

$$\sum_{j=1}^{n} b_j I_j, \qquad \text{where } b_j=\frac{1}{F'(\mu_j)}\int_{-1}^{1}\frac{\mu F(\mu)}{\mu-\mu_j}\,d\mu.$$

If, however, F is the Legendre function, we have in fact that these two sums are precisely the same, since

$$b_j-a_j\mu_j=\frac{1}{F'(\mu_j)}\int_{-1}^{1}\frac{\mu F(\mu)}{\mu-\mu_j}\,d\mu-\frac{1}{F'(\mu_j)}\int_{-1}^{1}\frac{\mu_j F(\mu)}{\mu-\mu_j}\,d\mu$$

$$=\frac{1}{P'(\mu_j)}\int_{-1}^{1}P(\mu)\,d\mu=0,$$

and the same holds for the integral

$$\int_{-1}^{1} I(\tau,\mu)\mu^2\, d\mu$$

which occurs in the Rayleigh-scattering case, provided the order of P is greater than one. The Newton method, of course, does not yield this relation.

With these considerations in mind, a new function was tried, namely

$$\frac{d^n}{d\mu^n}[\mu(1-\mu)]^n \quad \text{for} \quad 0<\mu<1,$$

and the same function with the sign of μ changed for $-1<\mu<0$. This function, which is simply the Legendre polynomial transformed into the half-interval, has zeros at neither zero nor unity, and both the zeros and the weights are symmetrically distributed in the four quarters of the interval $(-1, 1)$. In odd

orders, $\mu=\frac{1}{2}$ is a zero, and the weights in even orders are half the Gaussian weights in half the order. We are in fact using the Gauss formula twice, just as is often done with the simple formulae of numerical integration to secure greater accuracy. It will thus be seen that the calculations are considerably simplified, and exactness in the radiative transfer problem is enhanced. A few preliminary integrations of functions such as $\sin^4 \pi x/2$, $\tan \pi x/4$ and e^x between 0 and 1 gave excellent results, and the astrophysical calculations were then begun with every hope of success. Following Kourganoff, the q-function and the limb-darkening

$$I(0, \mu)=\tfrac{3}{4}F\left(\mu+Q+\sum_{\alpha=1}^{n-1}\frac{L_\alpha}{1+k_\alpha\mu}\right)$$

were calculated and the errors

$$\frac{\Delta(\tau+q)}{\tau+q} \text{ and } \frac{\Delta I_\mu}{I_\mu}$$

were found, the latter against the exact solution by Placzek (7). No more computation is needed than in either of the previous methods, and in fact (as indicated above) sometimes less. The objection against the Newton method, that the higher approximations are sometimes less accurate, holds against the new method (which may perhaps fittingly be called the double-Gauss method), but in a much more restricted range of τ, there being only one tabulated value where the third approximation is worse than the second. The correct evaluation of $\int I\mu d\mu$ and $\int I\mu^2 d\mu$, mentioned above, is naturally retained by the double-Gauss method.

The accompanying tables and graphs show the results obtained. For q, the second double-Gauss approximation is better than the third Newton and (except for $\tau>0{\cdot}7$) than the third Gauss. The new third approximation is much better than either of their fourths, especially for large τ. For limb-darkening, the second double-Gauss is better than the third Newton and the fourth Gauss. The new third approximation is far better than any other so far obtained, the relative error being less than 0·02 per cent over a great part of the range. The root-mean-square errors in each case were also computed, and emphasize the superiority of double-Gauss.

The accuracy of the various approximations for limb-darkening in the range $0<\mu<0{\cdot}1$ was also investigated, the maximum errors being Newton 2·33 per cent, Gauss 3·33 per cent and double-Gauss 1·38 per cent (second approximations). In the third approximation the new method again outdistanced the others.

The achievement of the greatest possible numerical accuracy, while not essential in the standard case, where the exact solution is known, is much more important in cases such as that of Rayleigh scattering, with the equation of transfer

$$\mu\frac{dI(\tau,\mu)}{d\tau}=I(\tau,\mu)-\frac{3}{16}\left[(3-\mu^2)\int_{-1}^{1}I(\tau,\mu')\,d\mu'+(3\mu^2-1)\int_{-1}^{1}I(\tau,\mu')\mu'^2 d\mu'\right],$$

and in more complicated situations, where exact solution is impossible. As shown above, the orthogonal property of the Legendre polynomials is a vital factor of success in these instances.

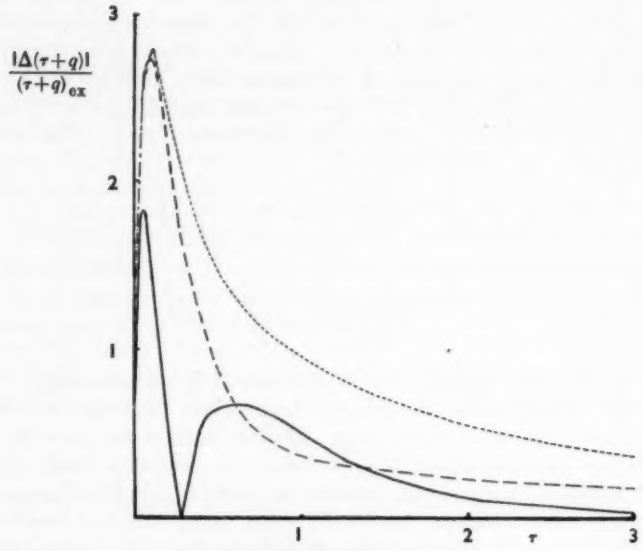


FIG. 1.—*Errors in source-function.*

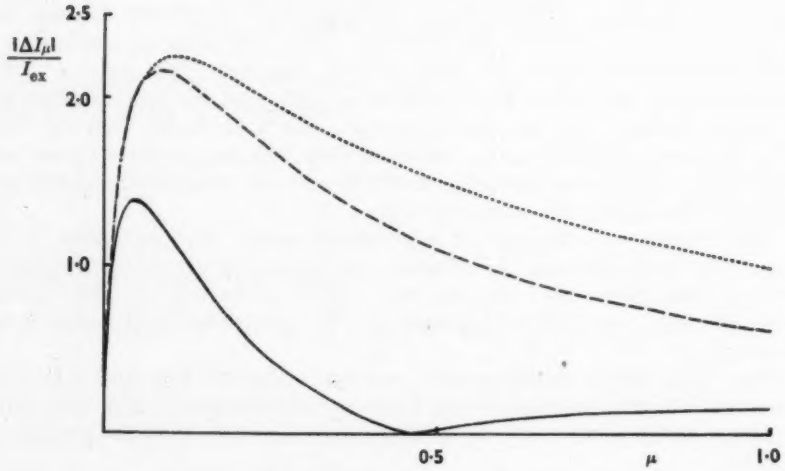


FIG. 2.—*Errors in limb-darkening.*

——— *Second double-Gauss approximation.*
– – – – *Third Gauss approximation.*
- - - - - - *Third Newton-Cotes approximation.*

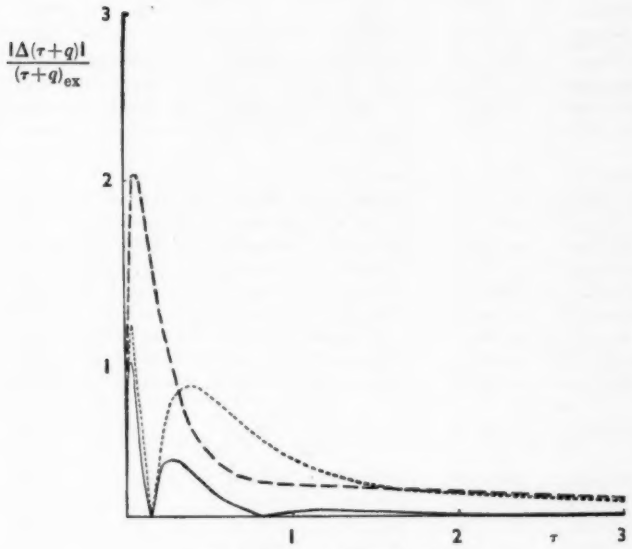


FIG. 3.—*Errors in source-function.*

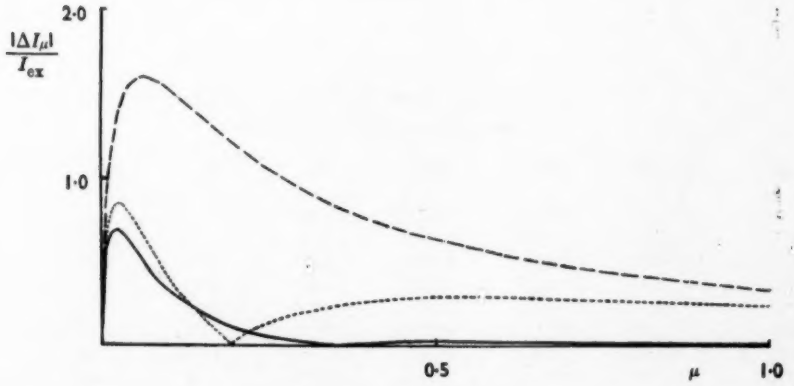


FIG. 4.—*Errors in limb-darkening.*

——— *Third double-Gauss approximation.*
– – – – *Fourth Gauss approximation.*
- - - - - - - *Fourth Newton-Cotes approximation.*

TABLE I*

Second approximations to $q(\tau)$

τ	q exact	q_{DG}	q_{NC}	q_G	Δ_{DG}	Δ_{NC}	Δ_G
0	0·577351	0·57735	0·57735	0·5774	0	0	0
0·01	0·588236	0·58191	0·58042	0·5796	−1·06	−1·31	−1·44
0·02	0·595391	0·58632	0·58342	0·5819	1·47	1·95	2·20
0·03	0·601242	0·59057	0·58639	0·5841	1·69	2·35	2·72
0·05	0·610758	0·59866	0·59217	0·5883	−1·83	−2·81	−3·40
0·1	0·627919	0·61657	0·60576	0·5982	−1·56	−3·04	−4·08
0·2	0·649550	0·64431	0·62965	0·6154	−0·62	2·34	4·02
0·3	0·663365	0·66393	0·64974	0·6295	+0·06	1·41	3·52
0·4	0·673090	0·67781	0·66664	0·6410	0·56	−0·60	2·99
0·5	0·680240	0·68762	0·68085	0·6505	0·63	+0·05	2·52
0·6	0·685801	0·69456	0·69279	0·6583	0·68	0·54	2·14
0·7	0·690123	0·69947	0·70285	0·6647	0·67	0·92	1·83
0·8	0·693534	0·70294	0·71130	0·6699	0·63	1·19	1·58
0·9	0·696276	0·70540	0·71840	0·6743	0·57	1·38	1·38
1·0	0·698540	0·70713	0·72438	0·6778	+0·51	+1·52	−1·22
1·2	0·701899	0·70923	0·73363	0·6831	+0·38	+1·67	−0·99
1·5	0·705130	0·71058	0·74269	0·6879	+0·25	+1·70	−0·78
2·0	0·707916	0·71119	0·75039	0·6918	0·12	1·57	0·60
2·5	0·709191	0·71130	0·75363	0·6931	0·07	1·38	0·50
3·0	0·709806	0·71132	0·75499	0·6937	0·04	1·22	0·43
3·5	0·710120	0·71132	0·75557	0·6939	0·03	1·08	0·39
4·0	0·710270	0·71132	0·75581	0·6940	0·02	0·97	0·35
5·0	0·710398	0·71132	0·75595	0·6940	+0·02	+0·80	−0·29
∞	0·710446	0·71132	0·75598	0·6940	0	0	0
	Root-mean-square error				0·79	1·54	2·06

* In each table, Δ denotes the relative error per cent.

TABLE IA

Second approximations to $I(0, \mu)$

μ	$\frac{4I_\mu}{3F}$: exact	DG	NC	G	Δ_{DG}	Δ_{NC}	Δ_G
0	0·43301	0·43301	0·4330	0·4330	0	0	0
0·1	0·54012	0·53386	0·5278	0·5224	−1·16	−2·28	−3·28
0·2	0·62802	0·62413	0·6175	0·6078	0·62	1·68	3·22
0·3	0·71123	0·70921	0·7038	0·6905	0·28	1·04	2·92
0·4	0·79210	0·79137	0·7878	0·7716	−0·08	0·54	2·59
0·5	0·87156	0·87171	0·8702	0·8515	+0·02	−0·16	2·30
0·6	0·95009	0·95085	0·9513	0·9304	0·08	+0·13	2·07
0·7	1·02796	1·02915	1·0314	1·0088	0·12	0·34	1·86
0·8	1·10535	1·10684	1·1108	1·0866	0·13	0·50	1·70
0·9	1·18238	1·18408	1·1896	1·1640	0·14	0·61	1·55
1·0	1·25912	1·26098	1·2679	1·2411	+0·15	+0·70	−1·43
	Root-mean-square error				0·44	1·03	2·38
	Maximum error				−1·38	−2·33	−3·33
	Value of μ for maximum error				0·04	0·08	0·13

TABLE II

Third and fourth approximations to $q(\tau)$

		Third			Third : Fourth				
τ	q_{exact}	q_{DG}	q_{NC}	q_{G}	Δ_{DG}	Δ_{NC}	Δ_{G}	Δ_{NC}	Δ_{G}
0	0·577351	0·57735	0·57735	0·5774	0	0	0	0	0
0·01	0·588236	0·58378	0·58092	0·5809	0·74	−1·22	−1·23	−0·84	−1·10
0·02	0·595391	0·58979	0·58436	0·5843	0·91	1·79	1·81	1·07	1·55
0·03	0·601242	0·59542	0·58767	0·5875	0·92	2·15	2·17	1·13	1·81
0·05	0·610758	0·60560	0·59395	0·5938	−0·78	−2·54	−2·57	−1·04	−2·02
0·1	0·627919	0·62593	0·60777	0·6080	−0·27	−2·77	−2·74	−0·49	−1·93
0·2	0·649550	0·65181	0·62901	0·6307	+0·27	2·42	2·22	+0·37	1·29
0·3	0·663365	0·66660	0·64404	0·6477	0·33	2·01	1·63	0·71	0·80
0·4	0·673090	0·67587	0·65487	0·6603	0·26	1·70	1·19	0·77	0·51
0·5	0·680240	0·68219	0·66283	0·6698	0·16	1·48	0·89	0·72	0·36
0·6	0·685801	0·68686	0·66879	0·6770	0·08	1·32	0·69	0·63	0·27
0·7	0·690123	0·69051	0·67337	0·6824	+0·028	1·20	0·55	0·54	0·22
0·8	0·693534	0·69349	0·67695	0·6866	−0·003	1·11	0·46	0·45	0·20
0·9	0·696276	0·69597	0·67982	0·6898	0·019	1·03	0·41	0·38	0·19
1·0	0·698540	0·69806	0·68217	0·6924	−0·028	−0·96	−0·36	+0·33	−0·18
1·2	0·701899	0·70136	0·68575	0·6959	−0·029	−0·85	−0·32	+0·24	−0·17
1·5	0·705130	0·70473	0·68940	0·6990	−0·018	−0·71	−0·28	+0·17	−0·16
2·0	0·707916	0·70784	0·69300	0·7016	−0·003	0·55	0·23	0·12	0·13
2·5	0·709191	0·70929	0·69495	0·7025	+0·003	0·44	0·21	0·10	0·11
3·0	0·709806	0·70997	0·69604	0·7033	0·004	0·37	0·18	0·08	0·09
3·5	0·710120	0·71029	0·69664	0·7036	0·004	0·32	0·16	0·08	0·08
4·0	0·710270	0·71044	0·69698	0·7037	0·003	0·28	0·14	0·07	0·07
5·0	0·710398	0·71054	0·69728	0·7038	+0·002	−0·23	−0·12	+0·06	−0·06
∞	0·710446	0·71057	0·69742	0·7039	0	0	0	0	0
			Root-mean-square error		0·36	1·40	1·21	0·56	0·86

TABLE IIA

Third and fourth approximations to $I(0, \mu)$

		Third			Third : Fourth				
μ	$\frac{4I_\mu}{3F}$: exact	DG	NC	G	Δ_{DG}	Δ_{NC}	Δ_{G}	Δ_{NC}	Δ_{G}
0	0·43301	0·43301	0·4330	0·4330	0	0	0	0	0
0·1	0·54012	0·53833	0·5280	0·5285	−0·33	−2·24	−2·15	−0·41	−1·52
0·2	0·62802	0·62740	0·6149	0·6164	0·10	2·09	1·85	+0·02	1·20
0·3	0·71125	0·71112	0·6979	0·7003	−0·02	1·87	1·54	0·19	0·93
0·4	0·79210	0·79218	0·7789	0·7819	+0·0101	1·67	1·29	0·26	0·76
0·5	0·87156	0·87170	0·8584	0·8620	0·0161	1·51	1·10	0·29	0·64
0·6	0·95009	0·95027	0·9371	0·9410	0·0189	1·37	0·96	0·29	0·55
0·7	1·02796	1·02816	1·0150	1·0193	0·0184	1·26	0·84	0·28	0·47
0·8	1·10535	1·10555	1·0925	1·0970	0·0172	1·16	0·76	0·27	0·42
0·9	1·18238	1·18257	1·1697	1·1743	0·0169	1·07	0·68	0·26	0·38
1·0	1·25912	1·25930	1·2465	1·2513	+0·0151	−1·00	−0·62	+0·24	−0·34
	Root-mean-square error				0·11	1·58	1·28	0·27	0·81
	Maximum error				−0·69	−2·24	−2·16	−0·84	−1·60
	Value of μ for maximum error				0·02	0·11	0·08	0·03	0·07

My thanks are due to Professor H. H. Plaskett for his help and encouragement, and to Mr T. D. Kinman who drew the graphs. I am also grateful to the Master and Fellows of Balliol College and to the Department of Scientific and Industrial Research for awards, during the tenure of which the work has been carried out.

University Observatory, Oxford:
1951 *March* 9.

References

(1a) S. Chandrasekhar, *Ap. J.*, **100**, 76, 1944.
(1b) S. Chandrasekhar, *Ap. J.*, **100** onwards.
(2) C. F. Gauss, *Coll. Works*, **3**, 163, 1816.
(3) Z. Kopal, *Ap. J.*, **104**, 74, 1946.
(4a) V. Kourganoff, *Ann. d'Ap.*, **12**, 169, 1949.
(4b) V. Kourganoff, *Ann. d'Ap.*, **12**, 247, 1949.
(5) C. Mark, *Phys. Rev.*, **72**, 558, 1947.
(6) I. Newton, *Principia*, **3**, xl, 5, 1687.
(7) G. Placzek, *Phys. Rev.*, **72**, 556, 1947.
(8) K. Schwarzschild, *Göttinger Nachrichten*, 41, 1906.
(9) G. C. Wick, *Zeits. f. Phys.*, **121**, 702, 1943.

AN EYEPIECE INTERFEROMETER

W. S. Finsen

(Received 1951 April 5)

Summary

An eyepiece interferometer is described, intended primarily for the observation of close double stars. Some of the more significant features of the instrument are :—

1. The interferometer unit is complete in itself, and only requires a collimator unit to adapt it to any telescope.
2. It is of convenient size, with an over-all length of approximately 5 inches.
3. As the slits are located at a pupil in a collimated beam the instrument is not sensitive to reasonable displacements of the star from the axis, nor does the slit constant depend critically on the location of the slit plane.
4. It may be "twirled" very rapidly between finger and thumb, so that a scanning procedure may be employed, of great value in eliminating spurious minima due to poor seeing.
5. The slit separation may be varied with the touch of a finger.
6. The star may be examined with full aperture at will.

---

1. *Introduction.*—In 1933 I improvised an experimental double-star interferometer of the Anderson type (**1**) for the measurement of a short list of close pairs with the 26½-inch refractor of the Union Observatory (**2**). Though the results so obtained were more than encouraging (**3**), the tedious operation of interchanging it with the filar micrometer at frequent intervals prevented its regular use. The obvious requirement was a combined instrument so that either micrometer or interferometer could be used at will with the minimum delay. An additional disadvantage of this interferometer was the cumbersome procedure for changing the slits and the necessity, in the absence of cross-threads to define the centre of the field, for continual checking of the projection of the slits on the object glass.

In 1948 an aluminium extension tube was fitted to the micrometer, directed towards the object glass and provided with a dowelled stage for the reception of slit diaphragms. A convenient hand-hole in the telescope tube gave access to the stage. As an afterthought, provision was made for varying the slit separation by means of a Bowden cable operated from the eye end, the construction of the micrometer precluding the use of a more positive method of control.

Though the interferometer could now be put into use at a moment's notice, there was still room for improvement. The Bowden cable limits rotation in position angle to one full revolution, and the micrometer itself is not suitable for the rapid rotation or "twirling" so desirable in the interferometric observation of double stars. The problem had, in fact, been narrowed down to the design of an external attachment to the micrometer capable of free and rapid rotation independently of it.

The solution becomes obvious on realization of the fact that the interferometer slits may be placed at any point in the optical train. The ideal location in front

of the object glass is ruled out for practical reasons. In the conventional Anderson interferometer the slits intersect the converging cone of rays at some point between the object glass and its focus. This arrangement is very sensitive to errors in the assumed location of the slit plane relative to the focal plane, and to displacements of the star from the collimation axis. Finally, there is no theoretical objection to placing the slits in the beam of parallel rays between the eyepiece and the eye itself. This is, in fact, the analogue of the ideal position in front of the object glass and possesses comparable advantages. In practice, however, serious difficulties would arise in registering the slits in the minute exit pupil resulting from the necessarily high magnifying power. But the same effect is easily achieved by placing the slits at the larger exit pupil yielded by a low-power eyepiece, using a small secondary telescope to give the desired final magnification.

As the slits are located at a pupil, the instrument will be insensitive to reasonable axial displacement of the image. It is assumed that collimation is obtained with the use of a converging lens; with a diverging lens (Galilean system) there is no real exit pupil at which to place the slits, and this advantage is lost.

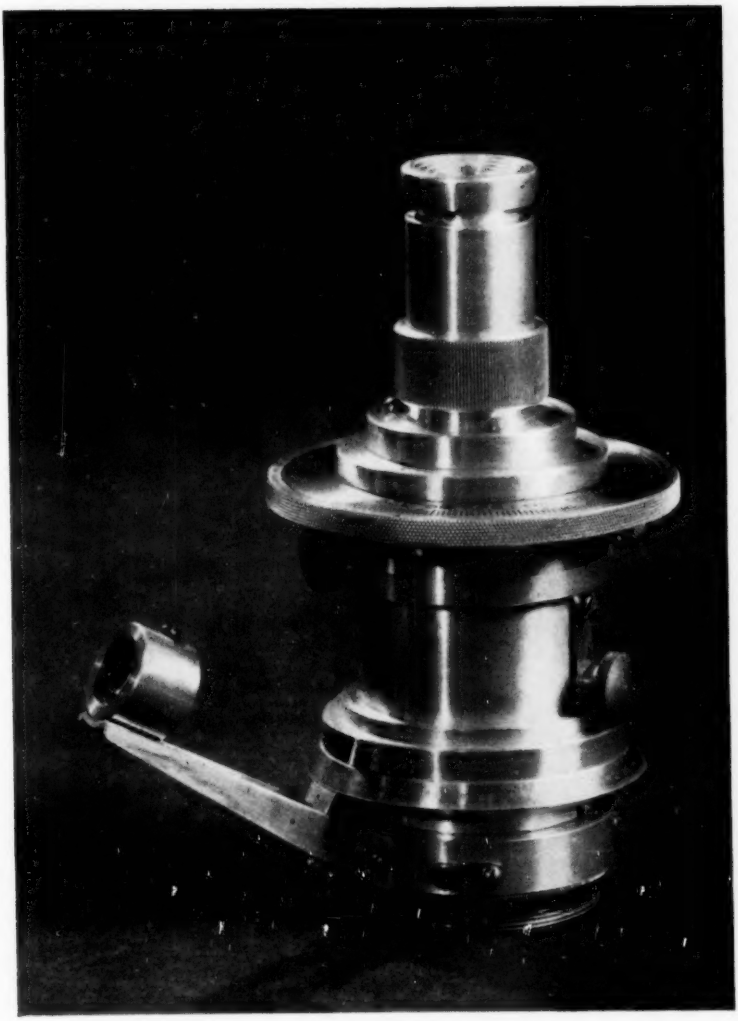
A low-power first eyepiece or collimator lens yields a large exit pupil, which is of advantage for minimizing the effect of errors of construction or register of the slits. But the exit pupil or eye-point is correspondingly more distant from the focal plane of the principal object glass; furthermore, a correspondingly higher power will be required of the secondary telescope. It is necessary, therefore, to effect a compromise in order to keep the dimensions of the instrument within reasonable limits without demanding impracticably small tolerances in the construction of the slit mechanism.

In 1950 September I completed an eyepiece interferometer of this type (Plate 5). Since then it has been in regular use, performing under practical conditions with entire satisfaction.

It is convenient to consider the whole instrument as consisting of two distinct units: a collimator and an interferometer.

2. *The collimator.*—This is shown at A in Plate 6. The collimator lens is an achromatic doublet of 37½ mm. focal length, mounted in a rigid housing which may be screwed into the eyepiece adapter socket of the Repsold micrometer. The flint component faces the incident light, and provision is made for accurate focusing, centring and squaring-on. It yields an exit pupil 2·31 mm. in diameter (corresponding to a power of 291) situated 4·1 mm. outside the housing. A sleeve D rotates freely but without shake about the axis defined by the optical centre of the collimator lens and the centre of the exit pupil. This collimation must be carried out with high accuracy, but the procedure is quite simple. The properly squared-on lens is first focused by examining the image of the micrometer threads with a telescope focused for infinity. It is then centred with respect to the axis of rotation by clamping the rotating sleeve in a vee-block, the same telescope being used to examine the image of an object placed at the anterior focus of the rotated collimator lens. The collimator housing is then replaced on the micrometer, and the exit pupil centred on the axis of rotation by means of three levelling screws not shown in the photograph. Even with a stable instrument this adjustment should be checked from time to time; a convenient way of doing this will be described later.

A ring B, carrying an index and reader for position-angle determination, slips over the rotating sleeve and is clamped to the non-rotating base of the housing.



W. S. Finsen, *An eyepiece interferometer.*

W. S Finsen, An eyepiece interferometer.

3. ***The interferometer.***—The interferometer C consists of a miniature telescope, focused for infinity, with an achromatic object glass of 35 mm. focal length. Three pre-focused eyepieces are available, with effective focal lengths of 16·4, 13·0 and 8·7 mm., yielding final powers of 620, 770 and 1170. All optical components, including the collimator lens, have been "bloomed".

The slit mechanism is located immediately in front of the miniature object glass. Two slit plates E slide freely but without play in a pair of grooves. Springs F press a stud on each plate against opposite sides of a wedge attached to the plate G which moves in a slide at right angles to the direction of movement of the slit plates. A stud on the wedge plate engages in turn in an eccentric groove cut in the face of a graduated dial H which rotates about the axis of the instrument. This gives a large reduction factor, a half-revolution of the dial corresponding to a linear displacement of about 1·14 mm. in each slit plate.

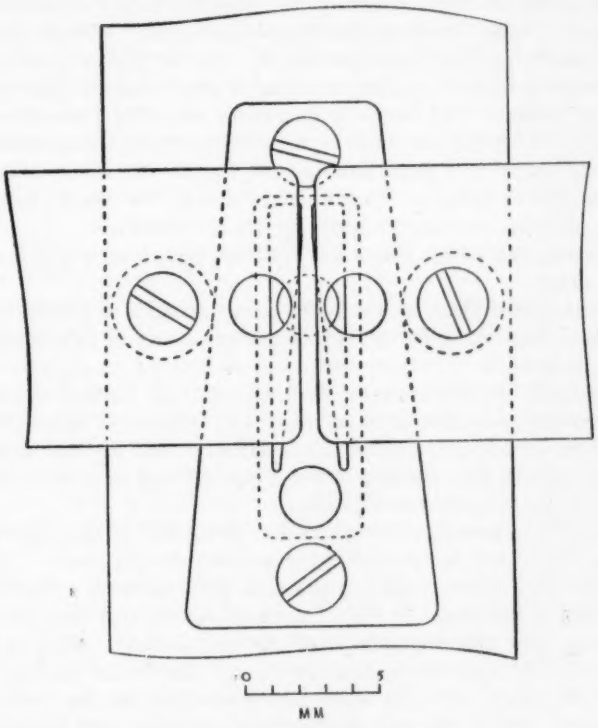


FIG. 1.—*Detail of slit mechanism.*

The slit apertures are cut very close to the opposing edges of the slit plates. The variable gap between the plates is optically closed by a second masking wedge formed by two triangular apertures in the register wedge mentioned earlier. At one extreme position of the dial the slit plates are withdrawn entirely from the optical train and a circular aperture in the register wedge permits the examination of a star with full aperture.

A tubular extension I of the interferometer carries a position circle and is a snug fit on the rotating sleeve of the collimator, where it is held in place by a spring catch J engaging in a groove.

It goes without saying that the slit mechanism must be very accurately constructed. Tolerances not exceeding 0·001 inch are desirable. It is equally obvious that the collimator lens and its exit pupil should be centred to the same order of accuracy. Strictly speaking, the same high degree of accuracy is perhaps not necessary in the case of the remaining optical components, but I have nevertheless found it well worth while, even in the case of the eyepieces, in order to prevent wandering of the image on rotating the instrument. This ensures also that the centre of the apparent field will coincide with the collimation axis and therefore (assuming a well-adjusted principal object glass) with the point of best fringe visibility.

4. *The slits.*—Two interchangeable sets of slit plates are available. In one pair the apertures are circular, and of the same size as the collimator exit pupil. The resultant "slits" are therefore lune-shaped, i.e. formed of two equal segments joined by their common chord. As the plates move inward their width increases with a consequent decrease of inter-centroid distance and hence of resolving power. The result is maximum resolving power for given light transmission. This has proved to be a very convenient arrangement—when the telescope is directed to a particular star and the slits are slowly opened till there is sufficient light to make the fringes clearly visible, one has the assurance of the maximum resolving power practicable in the circumstances.

The second pair of slit plates has rectangular apertures yielding "slits" of segmental shape.

It is somewhat difficult to assess the relative merits of lune-shaped and segmental slits. The lune may be regarded as composed of an infinite number of elementary rectangles of varying width but all centred on the axis of symmetry of the lune itself. In other words, the inter-centroid distance of corresponding pairs of elementary rectangles is constant. In the case of rectangular slits it is known that the slit constant depends on the inter-centroid distance alone, providing attention is paid to the visibility of the fringe pattern as a whole (**4**). This is therefore true for lune-shaped slits also.

Segmental slits may also be regarded as composed of elementary rectangles, but in this case both width and inter-centroid distance vary. The rigorous computation of the fringe spacing presents an intractable problem. When I first used slits of this shape in 1933 I assumed simply that they were equivalent to rectangular slits with the same inter-centroid distance. In this I have been followed by Wilson (**5**), who further points out that in the limiting case (never realized in practice), when the segments are semi-circles, this procedure gives a "slit constant" only 3·4 per cent smaller than the theoretical resolving power of an unobscured circular object glass. For infinitely narrow segments the error is, of course, zero. In computing the slit constants for the segmental slits I have made an adjustment for this error based on linear interpolation.

Apart from the slight uncertainty in the value of the slit constant, segmental slits seem to have some theoretical advantage over lune-shaped slits. For the same inter-centroid distance they transmit about 6 per cent more light, and the diffraction pattern is clearly narrower and probably longer. This was confirmed in a laboratory test. The shorter fringes were apparently brighter and easier to

see, and there was noticeably less contrast as regards both shape and brightness between central and extreme fringes. As a result both minima and "zig-zag effects" (2) were decidedly easier to recognize. A similar test was made under practical conditions at the telescope; variable seeing prevented a clear-cut decision but the conclusion was, on the whole, in favour of segmental slits.

With the present instrument used in conjunction with the 26½-inch refractor the slit constant can be varied between 0″·087 and 0″·132 with lune slits and between 0″·087 and 0″·143 with segmental slits. In both cases the smaller values correspond to infinitely narrow slits and are not realizable in practice.

A mask with a circular aperture of the same size as the exit pupil formed by the collimator can be placed immediately in front of the slit apertures. This is of doubtful value with a well-adjusted instrument and is usually dispensed with. It has the disadvantage of stopping-down the instrument as regards visibility of the micrometer threads.

5. *Adjustment of the interferometer.*—It is assumed that the optical components have been centred relative to the rotation axis as described earlier. The collimator exit pupil must now be centred, at the same time checking its registration with the slits. As this operation should be repeated from time to time, a special attachment has been provided for the purpose. This consists of an extension tube fitting over the eyepiece, with the help of which one of the remaining eyepieces can be focused on the final exit pupil. The slits are thus seen projected on the principal object glass, making it easy to centre the collimator exit pupil by means of the levelling screws mentioned earlier, at the same time checking the centring of the slits. (This simple procedure is useful also for checking the slit projection of Anderson interferometers.)

The position-circle zero is easily determined by directing the telescope at a bright star with narrow slits and aligning the elongated diffraction pattern with a micrometer thread.

With a well-adjusted instrument the fringes should have maximum visibility in all position angles at the centre of the field. Near the edge of the field this will in general no longer be true. When the fringes are radial their visibility will be but little impaired, but on rotating through 90° they will become more or less invisible. The effect is indistinguishable from genuine duplicity, but this erroneous conclusion is easily guarded against by keeping the star in the centre of the field.

It may happen, however, that the position of maximum fringe visibility does not coincide with the centre of the field. The effect is noticed with both the Anderson interferometer and the present instrument. On the assumption that the star is zenithal (for atmospheric dispersion has the same result) and that faulty adjustment of the interferometer has been eliminated as a possible cause, the most probable explanation is relative decentring of the object glass components, resulting in a prismatic effect. The fault is not cured by a reasonable squaring-on adjustment of the object glass, nor by moving the eyepiece laterally; the fringes remain invisible unless the direction of vision is more or less oblique—the pencil of rays passing through the eyepiece is thereby decentred so that a compensating prismatic component is introduced. If the star is placed centrally in the field it is easy to determine the sense in which the components of the object glass are decentred by placing a prism of small refracting angle between the eyepiece and the eye and noting the orientation that gives improved fringe visibility. The

object glass clearly requires prismatic correction in the sense indicated by the prism, making proper allowance for any ray inversion.

It may be remarked in passing that this procedure provides a very simple and sensitive method of adjusting the relative centring of object glass components. For this particular purpose the slits may be cut in cardboard and placed in the ideal position in front of the object glass, thus eliminating other extraneous factors.

It is also of interest to note that it is fallacious at large zenith distances to compare the fringe visibility of a suspected double with that of a neighbouring star as control, unless the precaution is taken of examining both stars at exactly the same point in the field. Fringes invisible at the centre of the field as the result of atmospheric dispersion may be readily visible nearer the edge with oblique vision. But the deliberate use of this fact to compensate for atmospheric dispersion cannot be recommended; it is too uncertain and it introduces the possibility of other off-axis errors (especially in the case of the Anderson interferometer). The safest procedure is the introduction of variable dispersion between the eyepiece and the eye by means of a rotary or Risley prism (**6**).

Union Observatory,
Johannesburg:
1951 *March* 27.

References

(1) J. A. Anderson, *Ap. J.*, **51**, 263, 1920.
(2) W. S. Finsen, *U.O.C.*, No. 90, 1933.
(3) P. Baize, *J. des O.*, **32**, 107, 1949.
(4) H. Spencer Jones, *M.N.*, **82**, 533, 1922.
(5) R. H. Wilson, Jr., *Pub. Univ. Penn., Astr. series*, **6**, part 4, 8, 1941.
(6) W. S. Finsen, *Monthly Notes Astr. Soc. of South Africa*, **10**, 7, 1951.

PHOTOELECTRIC RECORDING OF STELLAR OCCULTATIONS

H. E. Butler

(Received 1951 February 26)

Summary

A photo-cell is mounted near the focus of a 15-inch reflecting telescope and the photo-current, corresponding to the intensity of starlight falling on it, is displayed as the vertical movement of one of the spots of a double beam oscilloscope. This deflection is recorded photographically on horizontally moving paper. Seconds marks obtained photoelectrically from the Observatory's Mean Time pendulum clock are shown on the second spot of the cathode ray tube. When a star is occulted at the dark limb of the Moon the change in deflection is clearly visible and its time is easily measured with respect to the clock marks to about $\frac{1}{100}$ of a second.

The same method is used to determine the clock errors, the B.B.C. time signals being displayed directly on the tube instead of the star. As a result, the clock error with respect to the time signals can also be determined to about $\frac{1}{100}$ of a second.

If several measures of clock error are made before and after the occultation, it is believed that the time of the occultation can be determined to within 1 or 2 hundredths of a second of the time defined by the time signals.

---

1. *Introduction.*—The visual method of timing the occultations of stars by the Moon gives a time which is not only late by the personal error of the observer, but which is, in addition, only accurate to about $\frac{1}{5}$ of a second. The observer, using a telescope of reasonable aperture, starts a stopwatch when he sees a star disappear (or when he sees it appear, whichever the case may be) and then compares his watch with a carefully rated clock.

Referring for a moment only to *disappearances* at the dark limb, which are the occultations most frequently observed, errors can arise in several ways:

(*a*) The observer will know the expected time of the occultation with an accuracy which will depend upon the geographical position of his observatory with respect to those observatories for which occultations are given in the Almanacs. This uncertainty may amount to as much as one or two minutes. The majority of observable occultations occur when the Moon is between 7 and 14 days old and when, as a result, the dark limb will be invisible. In most cases, therefore, the observer will have no warning of the occultation and will start to observe the star well before it is actually occulted, and his eye will be appreciably fatigued when the star eventually disappears. The suddenness of this disappearance is always surprising, and frequently one knows that up to half a second has elapsed between seeing the star disappear and starting the stopwatch.

In comparing the watch and the clock, since the seconds hand of the clock can be seen and the clock-ticks can be heard, the greater part of the personal error is avoided. Nevertheless, random errors of up to about $\frac{1}{10}$ of a second are still to be expected.

(*b*) The stopwatch itself is seldom a precision instrument. It may even have a small error over the few minutes that it will be running, but the biggest error

will usually be that it moves in jerks of $\frac{1}{5}$ of a second at a time, which, coupled with the difficulty of reading, introduces an error of up to $\frac{1}{5}$ of a second.

(*c*) The master clock will always have an error, and we must consider how accurately this error can be measured. Our own procedure is either to look directly at the clock and compare it with the B.B.C. time signals heard from a radio set in the same room, or to do the same thing using a stopwatch as an intermediary. In either case the error will be up to $\frac{1}{5}$ of a second. These errors can be reduced to some extent by making several readings before and after the occultation and graphing the result, but even so the final error will be of the order of $\frac{1}{10}$ of a second.

(*d*) If a chronograph is in use, (*b*) is eliminated and both (*a*) and (*c*) are unaltered. We introduce, however, the troubles that can arise when contacts are in use on an accurate clock and also introduce electrical and mechanical lags, etc., in the pens.

When observing *appearances* at the dark limb, the errors due to (*b*), (*c*) and (*d*) are unaltered, but that due to (*a*) is considerably increased.

2. *Method.*—It appears that the following method enables one to time an occultation with respect to the broadcast time signals with an accuracy of 1 or 2 hundredths of a second.

A 1P21 photomultiplier tube has been mounted about 2 inches behind the focus of the 15-inch reflector. There are changeable diaphragms at the focus. The output from the photo-cell is fed through an amplifier to give vertical displacements to one of the spots of the double-beam cathode-ray tube of an oscilloscope. A camera records these deflections on horizontally moving photographic paper, the spots themselves moving only vertically.

The second spot of the cathode-ray tube is given a deflection once a second by the Observatory's Mean Time pendulum clock. As some trouble had been experienced with the electrical contacts upsetting the rate of the clock, it was decided to use a photo-cell and a light beam, and so to impress seconds marks on the oscilloscope record without interfering in any way with the clock.

The arrangement is shown in Fig. 1.

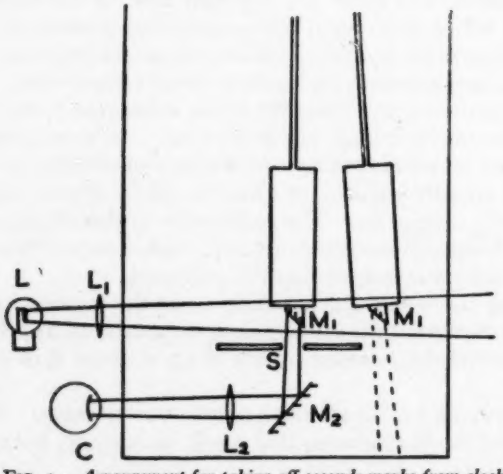


FIG. 1.—*Arrangement for taking off seconds marks from clock.*

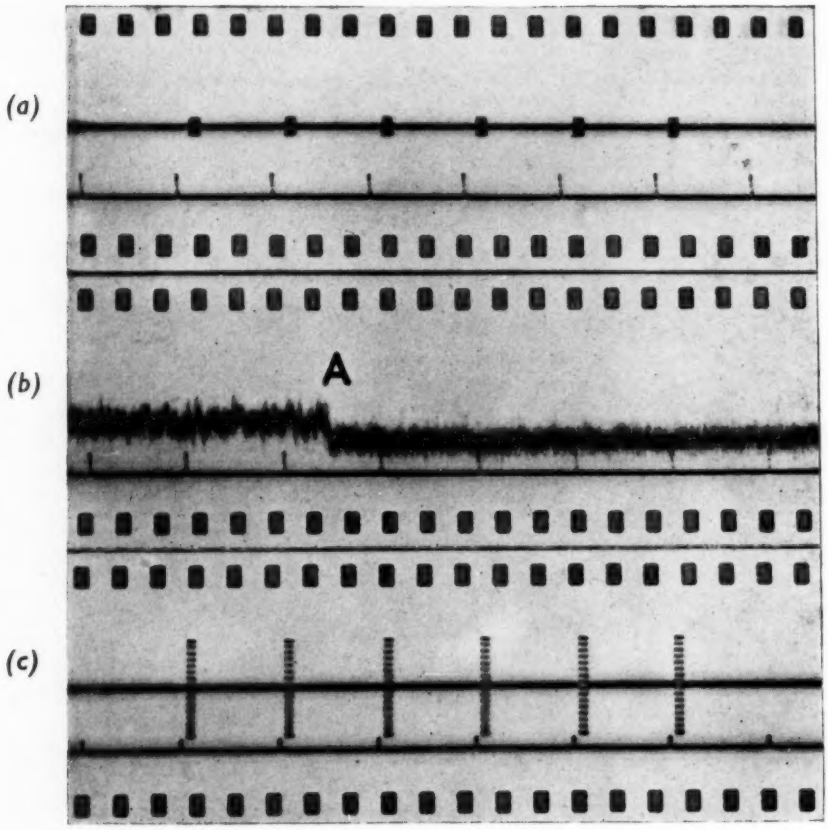


FIG. 2.—*Double beam oscillograms.*

Lower trace in each case shows seconds marks from Observatory clock.
Upper trace shows:

(*a*) *and* (*c*): *Broadcast six pips at* $19^h\,00^m$ *and* $22^h\,00^m$ *respectively.*

(*b*): *Trace of star γ Cancri, with occultation at point A.*

Consider the pendulum in its vertical position. A small plane mirror, 2 × 6 mm. (M_1) (actually a galvanometer mirror) was fixed to the point of the pendulum bob, with its longer side horizontal, and set at an angle of 45° to the horizontal. A small lamp L, with lens L_1, is mounted outside the glass clock-case and directs a beam of light horizontally so that the mirror M_1 is illuminated. Beneath M_1 is fixed a diaphragm with a slot S which is vertically under M_1 and the same size as the projected area of M_1. A portion of the original beam is reflected downwards by M_1, through the slot S and on to another mirror M_2. This mirror directs the light back horizontally through another lens L_2 and out through the case to fall on the cathode of a photomultiplier cell C.

As the pendulum swings, the cell is illuminated only as M_1 passes over S. The output current of the cell is fed through a cable to the telescope dome where it is displayed as the vertical deflection of the second spot of the oscilloscope. Each swing of the pendulum thus gives a short time-mark similar in shape to that of a chronograph. The duration of the mark is about 0·05 second. The lower traces of each of the three examples in Fig. 2 (Plate 7) show these marks when recorded with the relatively slow paper speed of 0·4 inch per second.

3. *Accuracy of rating of clock.*—Time signals are broadcast at most whole hours by easily received, medium wave-band transmitters of the B.B.C. from $05^h\,00^m$ to $23^h\,00^m$ and other times can be filled in, if necessary, with short-wave transmissions. The clock error is determined photographically at convenient intervals before and after the occultation, in the following way. Seconds are still shown on one spot of the oscilloscope, but the amplifier of the other spot, instead of being connected to the photo-cell on the telescope, is connected to the speaker terminals of a normal radio set. Sound is then displayed on the tube and, in particular, the six pips can be clearly seen and recorded. Examples are to be seen in the upper traces of (*a*) and (*c*) of Fig. 2 (Plate 7).

Fig. 3 shows clock errors obtained in this way on two occasions before and two occasions after an occultation, together with the best that could be obtained with

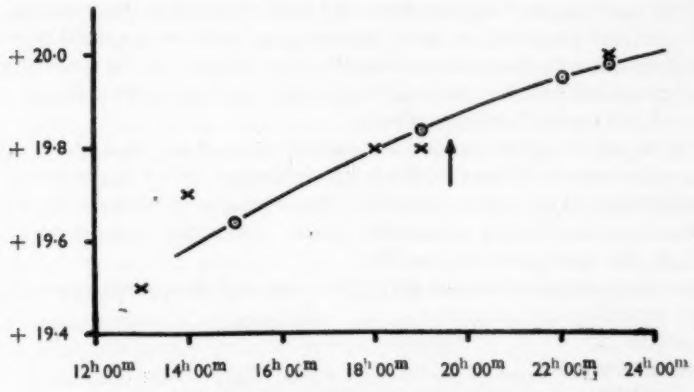


FIG. 3.—*Clock correction near time of occultation.*
Ordinates: Clock correction (seconds).
Abscissae: Time (1951 February 19).
The arrow indicates clock correction at time of occultation = + 19·87 seconds.
× : *Clock correction by stopwatch.*
○ : *Clock correction photographically.*

a stopwatch. Traces (*a*) and (*c*) of Fig. 2 (Plate 7) were the actual traces which were used to obtain the points at $19^h\,00^m$ and $22^h\,00^m$ respectively.

It will be seen that a very smooth curve can be drawn through the photographically determined points, the scatter being negligible. If we have a reasonably good clock and make frequent comparisons with the time signals, the time at any moment can in this way be determined to within $\frac{1}{100}$ of a second of the time as defined by the time signals themselves.

4. *Example of a recorded occultation.*—Fig. 2 (*b*) shows (Plate 7) the trace of the star γ Cancri, of magnitude 4·7, as it was occulted by the Moon (13·5 days old) on 1951 February 19.

Even with a focal aperture of 0·5 mm., considerable trouble was experienced with glare from the Moon, which was nearly full. As the Moon approached the star, the deflection of the spot steadily increased as a result of this scattered light.

The sudden decrease in deflection as the star was occulted is apparent as the downward movement of the trace at A. Prior to A (on the left), the scintillation of the star can just be seen on the trace, while after A only instrumental noise is present.

The speed of the recording paper was 0·4 inch a second and, needless to say, no idea of stellar diameter could be obtained. The break at A is sufficiently sharp for its time to be measured as $0{\cdot}47 \pm 0{\cdot}01$ second after the nearest seconds mark on the left. The occultation was also recorded by stopwatch in order to relate the fraction of a second so obtained to the appropriate whole minute and second. This enabled the time of the occultation to be written as $19^h\,37^m\,51^s{\cdot}47$ referred to the Observatory clock. Allowance for the clock error as read from the graph of Fig. 3 gave the final corrected time of the occultation, as obtained photographically, to be $19^h\,38^m\,11^s{\cdot}34$. Using a stopwatch both for the determination of clock error and for timing the occultation itself, gave a time of $19^h\,38^m\,11^s{\cdot}4$.

5. *Conclusions.*—The method described should give occultation times for disappearances at the dark limb as accurately as the broadcast signals warrant. It will be used on the brighter stars in future occultation programmes, both to obtain accurate times and to see what errors are present in normal observations. Actually, in order to attempt measures of stellar diameter at the same time, higher recording speeds will be used, although the resultant greater timing accuracy of the occultations is hardly justified.

It is hoped to try the method for *reappearances* at the dark limb. It will be necessary to centre the star in the focal diaphragm before it disappears behind the bright limb, and to rely on the telescope driving accurately for the whole time that the star is out of sight behind the Moon. Since this time can be as much as one hour, this may prove impossible.

The method is not suitable for appearances and disappearances at the bright limb.

Dunsink Observatory,
Co. Dublin:
1951 *February* 23.

MODELS FOR RED GIANT STARS

III. Models with a Single Discontinuity of Composition and Opacity due to Electron Scattering

C. M. Bondi and H. Bondi

(Received 1951 February 20)

Summary

Models consisting of a hydrogen envelope surrounding a core of higher molecular weight, with opacity throughout the star due to electron scattering, are examined. It is found that considerable (though only finite) extensions of radius can occur with a moderate fraction of the mass of the star in the envelope. The evolutionary tracks in the Hertzsprung–Russell diagram are examined, and good agreement with observation is found.

---

1. The present paper constitutes a further step in the systematic examination of the red giant problem which is the purpose of this series. In Paper I (**1**) the general features of the problem were discussed, and it was shown that no physically plausible law of opacity could lead to a homogeneous star that generates energy by the thermonuclear processes and possesses the large radius characteristic of red giants. Paper II (**2**) began the examination of stars possessing a single discontinuity of composition. The special opacity law $\kappa \sim Z(1+X)^{0.75}\rho^{0.75}T^{-3.5}$ was assumed to apply, and it was shown that large (and even mathematically infinite) radii could occur provided the change in composition was reasonably large and provided it took place in a certain part of the outer layers of the star. The model is then the Hoyle-Lyttleton model (**3, 4**) of a red giant, with a special law of opacity.

The corresponding problem for the case of opacity due to electron scattering forms the subject of the present paper. It seems likely that in actual red giant stars the law of opacity is intermediate between the case investigated in Paper II, where $\kappa \sim Z(1+X)^{0.75}\rho^{0.75}T^{-3.5}$, and the electron scattering case (**4**) in which $\kappa \sim (1+X)$. It may then be presumed that the structure of an inhomogeneous star with intermediate law of opacity is intermediate between the two cases. This will be especially likely if (as is the case) the investigation reveals that there are no really substantial differences between the two models, at least as far as the red giant problem is concerned. The problem of the structure of an inhomogeneous star with opacity due to electron scattering is, in any case, a subject of considerable interest.

2. The problem (first examined by Hoyle and Lyttleton (**4**)) may be stated as follows. The sphere of radius $r=r_2$ divides the star into two parts, each of which is homogeneous in composition. The abundances of hydrogen, helium, and other elements in the interior will be denoted by X_i, Y_i, Z_i, respectively, and the mean molecular weight there will be called μ_i. The corresponding quantities in the exterior region ($r>r_2$) will be called X_e, Y_e, Z_e and μ_e respectively. In accordance with the view that the difference in composition is due to the thermonuclear conversion of hydrogen into helium we shall assume that $Z_i=Z_e$. Since Z is taken to be very small in accordance with modern views, and is of no importance for the process of electron scattering, we shall put

$Z_i = Z_e = 0$. The material in the outer region is taken to be accreted interstellar matter in which the helium abundance is likely to be very small. Hence we shall put $Y_e = 0$ and therefore $X_e = 1$, $\mu_e = \frac{1}{2}$. The set of stellar models constructed therefore forms essentially a two-parameter series, the hydrogen abundance in the interior portion being one parameter (Y_i and μ_i being determined by it) and the mass ratio M_2/M_s being the other one.

The mode of energy generation in the large stars (in which alone electron opacity predominates) is almost certainly the Bethe cycle. This process is highly temperature sensitive. Therefore effectively all the energy generation will take place in the inner part of the star ($r < r_2$) provided that neither M_2/M_s nor X_i is too small. The first possibility concerns a case of little interest, the second may be shown to require X_i to be less than about 0·01. In the models to be considered the assumption will always be made that all the energy generation takes place in $r < r_2$, so that the model does not apply if X_i is less than 0·01. The case $X_i = 0$ will be considered but must be understood as a purely mathematical construction that can be approached fairly closely but not reached in reality.

Owing to the temperature sensitivity of the Bethe cycle the innermost part of the star will be in convective equilibrium. Denoting the radius of the convective core by r_1, it is clear that $r_1 < r_2$ since otherwise rapid mixing would occur. It will be assumed that all the energy generation occurs for $r < r_1$. This assumption greatly simplifies the treatment, but it also must be closely approached in any actual star.

Radiation pressure is not likely to be a factor of great importance in the models to be discussed in the paper, and will be neglected.

3. The outer part of the star ($r > r_2$) is either wholly or partly in radiative equilibrium according to the values of M_2/M_s and μ_i/μ_e. It can be shown that convective regions occur there only for rather small values of M_2/M_s, a case of little interest for the red giant problem. Just as in Paper II, the case in which there is a convective zone in the region $r > r_2$ will be excluded. The treatment is then very similar to that used in Paper II. The law of opacity in the radiative region in the present case is

$$\kappa = \kappa_0(1 + X) \tag{1}$$

where κ_0 is a constant, which in c.g.s. units has a value of about 0·19.

The energy transport equation is therefore

$$\frac{dT}{dr} = -\frac{3\kappa_0(1+X)L\rho}{16\pi acr^2T^3}. \tag{2}$$

We define h-variables as in the previous papers in this series (**1**, **2**), that is

$$S = \frac{4\pi Pr^4}{GM^2}, \quad Q = \frac{Pr}{GM\rho}, \quad N = 1 - \frac{d(\log\rho)}{d(\log P)}. \tag{3}$$

Then the hydrostatic condition is expressed by

$$\frac{S}{Q}\frac{dQ}{dS} = \frac{S+N-Q}{1+2S-4Q}, \tag{4}$$

while in the radiative region, the energy transport equation is by (2) and (3)

$$\frac{S}{N}\frac{dN}{dS} = \frac{S+1-4N}{1+2S-4Q}. \tag{5}$$

In the convective core $N = 0{\cdot}4$.

At the interface $r=r_2$, the quantities P, T, M and L are continuous, and hence S is continuous. Since P and T are both continuous, ρ/μ is continuous and hence from (2)

$$\frac{1}{\mu(1+X)}\frac{dT}{dr}$$

is continuous. If we denote the values of the discontinuous variables just outside the boundary $r=r_2$ by the suffix e and those just inside $r=r_2$ by the suffix i we then have

$$\frac{\rho_{2e}}{\rho_{2i}}=\frac{\mu_e}{\mu_i}\,;\quad \frac{Q_{2e}}{Q_{2i}}=\frac{\mu_i}{\mu_e}\,;\quad \frac{N_{2e}}{N_{2i}}=\frac{1+X_e}{1+X_i}. \tag{6}$$

The star may be divided into three zones as follows:

(*a*) $r_s\geqslant r>r_2$. An envelope in radiative equilibrium with composition $(X_e, Y_e, Z_e\,;\ \mu_e)$ (taken to be (1, 0, 0 ; $\frac{1}{2}$) respectively).

(*b*) $r_2\geqslant r>r_1$. An intermediate zone of composition $(X_i, Y_i, Z_i\,;\ \mu_i)$ in radiative equilibrium.

(*c*) $r_1\geqslant r\geqslant 0$. An inner convective zone having the same composition as zone (*b*).

Zone (*c*) always exists and so the range of cases arising depends on the relative extents of zones (*a*) and (*b*). In order that the arrangement shall be stable, we must have $\mu_e<\mu_i$. In the case in which zone (*a*) vanishes, the star is homogeneous and the corresponding solution will be referred to as the *standard solution*.

In zone (*b*) equations (4) and (5) apply and we have a second-order system. Only a one-parameter family of solutions of this system is of interest in zone (*b*) since at $S=S_1$ no change in composition occurs and consequently all the variables are continuous at $S=S_1$. We have to consider in (*b*) therefore only those solutions of (4) and (5) that fit smoothly on to the solution for the core. This is the Emden solution

$$Q=Q_E(S;0{\cdot}4),\ N=0{\cdot}4. \tag{7}$$

We therefore require to investigate the solutions of (4) and (5) for $r\geqslant r_1$, and this condition implies $S\leqslant S_1$. The value of S_1 may be taken as the parameter specifying the members of the family. A particular member of the family is the standard solution in which $S_1=S_1^*\simeq 1{\cdot}03$, which runs to the surface $S=Q=0$, $N=\frac{1}{4}$ (**5**). This standard solution will be written

$$Q=Q^*(S),\ N=N^*(S). \tag{8}$$

As in the problem solved in Paper II, the other members of the family lie close to the standard solution, provided that the value of N does not differ by a large amount from the standard value. These solutions diverge very rapidly from the standard once the N values deviate markedly. A higher approximation to the solutions can be obtained as follows. Consider

$$\frac{d}{dS}\log(NQ^4)=\frac{1+5S-4Q}{S(1+2S-4Q)} \tag{9}$$

obtained from (4) and (5). It will be seen that the rate of change with S of $\log(NQ^4)$ depends on Q and S only. It follows that if at any point $S=S_p$ of the zone (*b*) solution, $N-N^*$ is of the same order of magnitude as N^*, then at that point and for all $S>S_p$, Q is close to Q^*. By virtue of the first approximation the

right-hand side of (9) is very close to its standard values and hence a higher approximation is given by the relation

$$NQ^4 = N^*Q^{*4} \tag{10}$$

provided that $N-N^*$ does not differ by a large factor from N^*. Equation (10) holds within a few per cent if this condition is satisfied. No greater accuracy is likely to be required in zone (*b*).

In zone (*a*) the solutions also form a one-parameter family of solutions of (4) and (5) in the (S, Q) plane, since they must all run to the surface $S=Q=0$, $N=\frac{1}{4}$. Near the surface, these solutions are given approximately by

$$Q = AS^{1/4} + O(S^{5/4}), \quad N = \tfrac{1}{4} + O(S) \tag{11}$$

where A is a constant. A may be used to label the members of the family, which has precisely one member in common with the family (*b*), namely the standard solution with $A=A^*$. As in the problem discussed in Paper II, we are interested only in solutions with $A \geqslant A^*$.

The similarity with the problem in Paper II breaks down when we consider the curve described there as the spiral F. In the present case, the spiral point is the point $S=0$, $Q=N=\frac{1}{4}$. The spiral F, if it existed, would also start from this point. In fact no such solutions exist, and F may be considered to reduce to the point $S=0$, $Q=\frac{1}{4}$, $N=\frac{1}{4}$.

All members of the family with A greater than some value $\overline{A}$ ($\overline{A}<A^*$) run to the point $S=0$, $Q=N=\frac{1}{4}$ (*cf.* Paper I). Those with $A<\overline{A}$ run off to infinity. The curves fill the entire positive quadrant of the (S, Q) plane and there is one and only one member of the family through each point of this quadrant.

A definite value of N therefore corresponds to each point of the (S, Q) plane, namely the value taken by N at that point by the unique member of the family through the point. Curves of constant N can hence be drawn in the (S, Q) plane. Reference to Fig. 1* shows that the variation of N is slow and smooth.

As a solution is followed into the star from $S=Q=0$, $N=\frac{1}{4}$, the expression $dS/(1+2S-4Q)$ is always positive and so are N and S. Therefore, N will increase or decrease according as $S+1-4N$ is positive or negative. It follows that N cannot first reach any given value $N_p>\frac{1}{4}$ unless $S>4N_p-1$. Hence convective instability cannot first occur for $S<0{\cdot}6$. It will be shown later that large radii occur only for values of S_2 much smaller than $0{\cdot}1$, and hence the case we are discussing now is the one of interest.

Just as a definite value of N can be associated with each point of the plane, so also can a definite value of M/M_s and r/r_s be associated with each such point. These quantities are given by the quadratures

$$\log \frac{M_s}{M} = \int_0^S \frac{dS}{1+2S-4Q}, \tag{12}$$

$$\log \frac{r_s}{r} = \int_0^S \frac{Q\,dS}{S(1+2S-4Q)} = \int_0^Q \frac{dQ}{S+N-Q}. \tag{13}$$

The definite values are obtained by integration along the solution through the point. Thus curves of constant M/M_s and constant r/r_s can be drawn on the diagram representing the envelope solutions (Figs. 1, 2). The curves of constant M/M_s start on the S axis with M/M_s equal to $(1+2S)^{-1/2}$. They all run to the point $S=0$, $Q=\frac{1}{4}$. The curves of constant r/r_s start on the Q axis with $r/r_s=\frac{1}{4}-Q$ and run to infinity in the (S, Q) plane. The locus $r/r_s=0$ in

* The figures given in this paper summarize the results of eight integrations.

this problem consists only of the line $S=0$ for $Q \geqslant \frac{1}{4}$, and there is no curve corresponding to the spiral F. It follows that no stars of infinite radius can be constructed with this law of opacity since the equations (6) make the points on $S=0$ of no interest as points of fit.

In constructing a complete stellar model, the solution for zone (a) has to be fitted on to that for zone (b). At the fitting point, S, μQ and $N/(1+X)$ are continuous. The composition has been discussed in Section 2. As stated there, we shall take $Z_e=Z_i=0$ and $X_e=1$, $Y_e=0$. It follows that the maximum discontinuity possible occurs when $Y_i=1$ and $X_i=0$. This condition would be produced when all the hydrogen in the region $r<r_2$ had been converted into helium in the energy production processes.

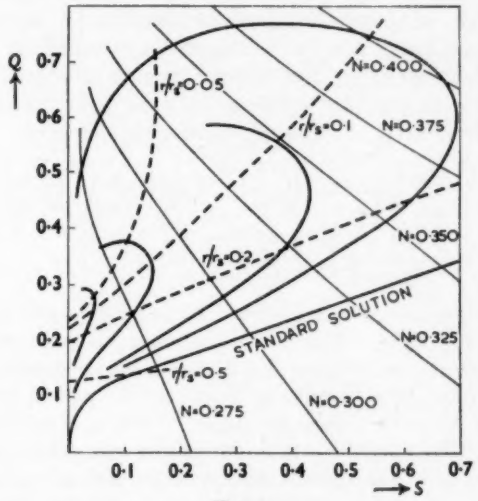


FIG. 1.

Thick lines : Solutions (A=0·218, 0·250, 0·270, 0·338, 0·429).
Thin lines : Curves of constant N.
Dashed lines : Curves of constant r/r_s.

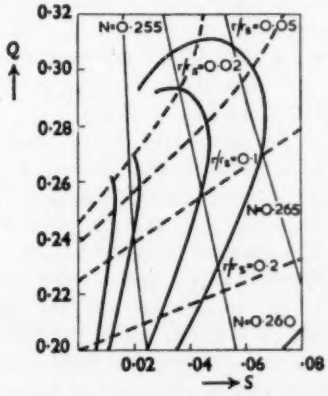


FIG. 1 (*a*).—*Detailed diagram of neighbourhood of $S=0$, $Q=0·25$ (solutions drawn for A=0·338, 0·429, 0·483, 0·593, 0·695).*

4. In the envelope, the solution will be given by one of the curves (*a*) already described. At $S=S_2$, both N and Q jump in accordance with equations (6) to one of the curves (*b*). For any S_2, the curves (*b*) relate Q_{2i} to N_{2i}. Therefore for any given change of composition, there corresponds to each value of S_2 a single set of values (Q_{2e}, N_{2e}), (Q_{2i}, N_{2i}), such that the points (S_2, Q_{2e}, N_{2e}) and (S_2, Q_{2i}, N_{2i}) lie respectively on a curve (*a*) and a curve (*b*), and such that equations (6) are satisfied. The locus (S_2, Q_{2e}) followed as S_2 varies will be termed the "fitting curve". The point at which the fitting curve intersects any curve $M_2/M_s=$constant, gives the correct value of S_2 and Q_{2e} for a model having a given change in composition and a given proportion $(M_s-M_2)/M_s$ of the mass in the envelope.

To construct the fitting curve accurately would involve performing a considerable number of integrations both of solutions (*a*) and of solutions (*b*). However, the fitting curve can be constructed approximately with very little calculation by making use of the fact that the zone (*b*) solution lies (as far as Q is concerned) close to the standard solution. We consider the ratio Q_{2e}/Q_2^* and assume (10) to hold exactly. Putting $X_e=1$ as usual, we have

$$\frac{Q_{2e}}{Q_2^*}=\frac{Q_{2e}}{Q_{2i}}\frac{Q_{2i}}{Q_2^*}=2\mu_i\left(\frac{N_2^*}{N_{2i}}\right)^{1/4}=2\mu_i\left(\frac{N_2^*}{N_{2e}}\right)^{1/4}\left(\frac{N_{2e}}{N_{2i}}\right)^{1/4}$$

$$=2\mu_i\left(\frac{2}{1+X_i}\right)^{1/4}\left(\frac{N_2^*}{N_{2e}}\right)^{1/4}. \tag{14}$$

Thus the fitting curve can be obtained by scaling up the standard solution by the factor on the right. This factor involves the ratio N_2^*/N_{2e}, which is determined completely only when the fitting curve is known, but is in fact always close to unity. The following limits can be placed on N_2^*/N_{2e}. The curves of constant N are very steep for small S, and as S tends to zero, N_2^*/N_{2e} approaches unity. It is found from computation that, provided no convective zone occurs for $r>r_2$, the curves of constant N are always sufficiently steep and sufficiently widely spaced for N_2^*/N_{2e} to exceed $\frac{1}{2}(1+X_i)$. Thus we have the inequality

$$2\mu_i \leqslant Q_{2e}/Q_2^* \leqslant 2\mu_i[\tfrac{1}{2}(1+X_i)]^{-1/4}, \tag{15}$$

Q_{2e}/Q_2^* therefore lying within fairly close bounds. If we take for a given S_2 a Q_{2e} satisfying (15) we can then find the value of N_2^*/N_{2e}. Using this we can find a second approximation to Q_{2e}/Q_2^*. This can be continued until a sufficiently accurate approximation to the fitting curve is obtained. The second approximation is generally sufficient.

If the fitting curve is drawn on the (S, Q) diagram for the envelopes displaying the curves $M_2/M_s=$constant (Fig. 2), then, as in Paper II, we can select the curve corresponding to a stellar model having an assigned fraction of the mass in the envelope. The point (S_2, Q_{2e}) is the point at which the fitting curve intersects the curve $M_2/M_s=$constant, and the values of N_{2e} and r_2/r_s may also be read off (Fig. 1). The quantities Q_{2i} and N_{2i} can be calculated from (6) and the intermediate solution (*b*) constructed if the standard solution is not considered to be sufficiently close to the actual solution.

5. It is of great importance to have some measure of the radius of a model so that one can compare the radius with that of the standard model. Such a measure must contain the information given by the Bethe law that the central temperature of all models considered must be approximately the same.

Consider the quantity

$$\log\left(\frac{\Re T_c}{\mu_i}\frac{r}{GM}\right) = \log Q + \int_S^\infty \frac{N\,dS}{S(1+2S-4Q)}, \tag{16}$$

the point (S, Q) being chosen to be a point of the solution where $r \leqslant r_2$. Since for $r \leqslant r_2$ we have Q and N approximately equal to their standard values, it follows that $\log Q$ and the integral on the right-hand side are close to their standard values. It follows that for stars of the same central temperature and central composition, the value of r_2/M_2 is close to its standard value.

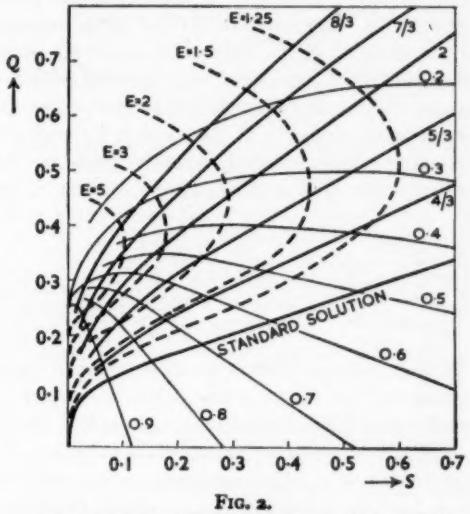


FIG. 2.

Thick lines : Fitting curves ($\mu_i/\mu_e = 1, \frac{4}{3}, \frac{5}{3}, 2, \frac{7}{3}, \frac{8}{3}$).
Thin lines : Curves of constant M/M_s.
Dashed lines : Curves of constant E.

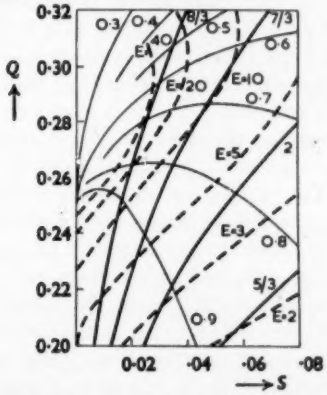


FIG. 2 (*a*).—*Detailed diagram of the neighbourhood of* $S=0$, $Q=0{\cdot}25$.

We now introduce a special form of (16), namely the invariant D used in both Papers I and II.

$$D = \frac{r_s}{GM_s}\frac{\Re T_c}{\mu_i}. \tag{17}$$

It again forms a convenient quantity to measure the radii of various models. In h-variables, we have

$$\log D = \int_0^{S_p} \frac{(Q-S)dS}{S(1+2S-4Q)} + \log Q_p + \int_{S_p}^{\infty} \frac{NdS}{S(1+2S-4Q)} \qquad (18)$$

where (S_p, Q_p) is an arbitrary point of the particular solution. If, as is convenient, we choose $(S_p, Q_p)=(S_2, Q_{2t})$, we know that the second and third terms in the expression on the right-hand side are almost the same for a non-homogeneous model as for the standard model. If D^* is the value of the invariant (17) for the standard model, it follows that the ratio D/D^* depends only on the changes in the first integral in the non-homogeneous model as compared with the standard model. This ratio is clearly of great importance, and is as before called the extension factor and denoted by E. It is the ratio of the radius of any model to that of a homogeneous star of the same mass, central temperature and central composition. If the models we wish to compare do not have the same central composition, allowance can be made for this by multiplying E by an appropriate factor (the ratio of the central molecular weights).

We have

$$\log E \simeq \int_0^{S_s} \frac{(Q-S)dS}{S(1+2S-4Q)} - \int_0^{S_s} \frac{(Q^*-S)dS}{S(1+2S-4Q^*)} = \int_0^{S_s} \frac{(Q-Q^*)(1-2S)dS}{S(1+2S-4Q)(1+2S-4Q^*)}. \qquad (19)$$

We see, as in Paper II, that from consideration of the signs of the factors in the last integral, no larger values of E can occur to the right of $S=0{\cdot}5$ than occur to its left.

In practice, we calculate E most easily by comparing the ratio $Mr_s/M_s r$ at the point (S_2, Q_{2c}) with its value for the standard solution at $S=S_2$. In this way we compare the radius of a non-homogeneous star with the radius of the standard-solution star having the same mass, central temperature and central composition. There is one value of E corresponding to each point (S_2, Q_{2c}) so that curves of constant E can be drawn in the envelope diagram (Fig. 2). On the standard solution $E=1$, while E is infinite only at points on the Q-axis for $Q \geqslant \frac{1}{4}$. The curves of constant E start on the Q-axis with $E=(1-4Q)^{-1}$. For a definite value of the ratio μ_i/μ_e and a definite value of M_2/M_s the corresponding E value may be read off the diagram containing the fitting curve and the curves of constant M_2/M_s. The relevant E value is that corresponding to the point at which the curve representing the ratio M_2/M_s intersects the fitting curve.

Since the only points at which E is infinite lie on the Q-axis above $Q=\frac{1}{4}$, it follows that for no finite value of μ_i/μ_e can an infinite extension factor be obtained. For, finite values of μ_i/μ_e give rise to fitting curves which reach the Q-axis only at the origin. If we restrict ourselves to values of μ_i/μ_e not greater than $\frac{8}{3}$, that is, the case of pure hydrogen for $r>r_2$ and pure helium for $r<r_2$, then the maximum value of E obtained appears to be in the neighbourhood of 36. If for given μ_i/μ_e we plot E against M_2/M_s, we find that as M_2/M_s decreases, that is, as the mass in the envelope increases, E first increases, then reaches a maximum and then decreases again as further mass is added to the envelope (Fig. 3).

6. We now have to deal with the values of the luminosity for these non-homogeneous models. With the opacity law used here

$$\frac{NQ^4}{S} = \left(\frac{\mathfrak{R}}{\mu G}\right)^4 \frac{1}{(4\pi)^2} \frac{3K_0(1+X)L}{4acM^3} \qquad (20)$$

at all points.

Applying this at the surface, we have the mass-luminosity relation

$$\tfrac{1}{4}A^4 = \left(\frac{\mathfrak{R}}{\mu G}\right)^4 \frac{1}{(4\pi)^2} \frac{3K_0(1+X)L}{4acM^3}. \qquad (21)$$

With this opacity law, we see that the radius does not enter into the relation, so that the relation between mass and luminosity does not depend, as in the case discussed in Paper II, on the central temperature.

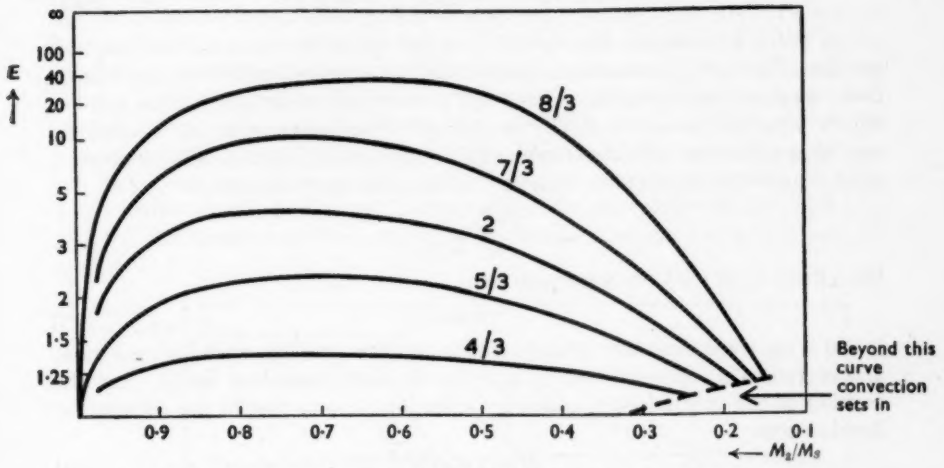


FIG. 3.—*Variation of extension factor E against* M_2/M_s *for various ratios* μ_2/μ_e. *Vertical scale:* $1-E^{-1/3}$.

By computation, it is found that a semi-empirical relation connects the luminosity of a non-homogeneous star with that of a certain homogeneous star. Equation (21) shows that assuming the same opacity law to apply to both stars and that their masses are equal, the luminosities are proportional to $A^4\mu_e^4(1+X_c)^{-1}$. If we take a homogeneous star with the same mean value of the composition as the non-homogeneous star, then it is found by computation that $A^4\mu_e^4(1+X_c)^{-1}$ is almost exactly the same in the two models. This property appears to hold for all non-homogeneous models with this opacity law, whether the extension factors are large or not.

7. In order to obtain a clearer view of the evolutionary significance of the theory discussed in this paper, the application of the theory to actual stars will now be considered.

It will be seen from Fig. 3 that appreciable values of the extension factor only occur if $0{\cdot}95 > M_2/M_s > 0{\cdot}4$. This implies that the discontinuity in composition occurs not at the surface of the convective core but somewhere in the outer layers of the star. This is a characteristic feature of the Hoyle-Lyttleton model of red giants. It can arise only if a well-mixed star that has converted a considerable fraction of its hydrogen into helium suddenly begins to accrete interstellar hydrogen at such a rate that it cannot be mixed in with the rest of the star. The required conditions are therefore that the speed of accretion should exceed the speed of mixing, which in turn must exceed the rate of helium formation (**6**). The currents recently discussed by Sweet (**7**) are presumably responsible for mixing the helium produced in the core throughout the original envelope of the star.

Assuming these conditions to be satisfied, the evolutionary sequence can now be traced.* A star of mass $\overline{M}$, radius $\bar{r}$ and luminosity $\overline{L}$, consisting of almost pure hydrogen, evolves through the process of helium formation. The rate of mixing is taken to be high enough to ensure a homogeneous composition. Then the new values of M, r and L are related to $\overline{M}$, $\bar{r}$, $\overline{L}$ by

$$M=\overline{M}, \quad L=\frac{32\mu^4}{1+X}\overline{L}, \quad r=2\mu\frac{\overline{T_c}}{\overline{T_c}}\bar{r}. \tag{22}$$

In order to consider the variation of the radius we have to consider the process of energy generation in some detail. For stars sufficiently massive to make electron opacity predominant the process of energy generation will be highly temperature-sensitive (Bethe cycle). The assumption of a constant central temperature will therefore be a fair approximation, and will be adopted since it greatly simplifies the analysis. Then since the invariant

$$\frac{\mathfrak{R}T_c}{\mu_i}\frac{r_s}{GM_s}$$

has a fixed value for a homogeneous star,

$$r=2\mu\bar{r}. \tag{23}$$

At some stage this slow development ceases to control the speed of evolution, and accretion is supposed now to become the most important factor. Let the fraction of the original mass added by accretion be ν so that in the subsequent development

$$M=(1+\nu)\overline{M}. \tag{24}$$

Since $\overline{M}$ is the mass of the part of the star having molecular weight μ_i the ratio $M_2/M_s=1/(1+\nu)$.

Since μ_c and X_c are taken to be constant throughout the subsequent development, the luminosity will vary according to

$$L\sim A^4M^3\sim A^4(1+\nu)^3. \tag{25}$$

But for very small ν the luminosity must be unchanged by the thin hydrogen envelope. Hence, by (22) and (14) (since, for very small S_2, $A/A^*=Q_{2c}/Q_2^*$)

$$\frac{L}{\overline{L}}=\left(\frac{A}{A^*}\right)^4(1+\nu)^3. \tag{26}$$

Also, for constant T_c and μ_i

$$r\sim EM\sim E(1+\nu).$$

Combining this with (22) we have by continuity at $\nu=0$

$$\frac{r}{\bar{r}}=2\mu_iE(1+\nu). \tag{27}$$

In order to construct evolutionary tracks in the Hertzsprung-Russell diagram it is necessary to evaluate the effective surface temperature T_{eff} in the various cases. Since $L\sim r^2T_{\text{eff}}^4$ this is easily derived. In the early stages of evolution

$$\frac{T_{\text{eff}}}{\overline{T}_{\text{eff}}}=(2\mu)^{1/2}\left(\frac{2}{1+X}\right)^{1/4}. \tag{28}$$

* The suffix s for surface values will be omitted in this section since M and r always refer to the mass and radius of the whole star.

After accretion has started to predominate we have

$$\frac{T_{\text{eff}}}{\overline{T}_{\text{eff}}} = \frac{A}{A^*}\,\frac{(1+\nu)^{1/4}}{(2\mu_i E)^{1/2}}. \tag{29}$$

The construction of the tracks is now a simple matter. The diagram (Fig. 4) represents absolute magnitude against the logarithm of the effective temperature. Starting from a definite luminosity $\overline{L}$ and temperature $\overline{T}_{\text{eff}}$ at $P_{1/2}$ the first slow development forms the straight line $P_{1/2}\,P_{4/3}$. When rapid accretion begins, the track turns sharply right. The point of the straight line $P_{1/2}\,P_{4/3}$ at which this development begins will vary according to the stage of evolution along $P_{1/2}\,P_{4/3}$ that had been reached when rapid accretion began, *i. e.* according to the value of μ_i. Evolutionary tracks PQRS are therefore drawn for 5 representative values of μ_i ($\frac{2}{3}$, $\frac{5}{6}$, 1, $\frac{7}{6}$, $\frac{4}{3}$). The last forms a limiting case that can be almost though not quite reached. For, as was stated before, when X_i drops below about 0·01 a considerable part of the energy production will take place in the outer zone of the star, and the model here used will no longer be applicable.

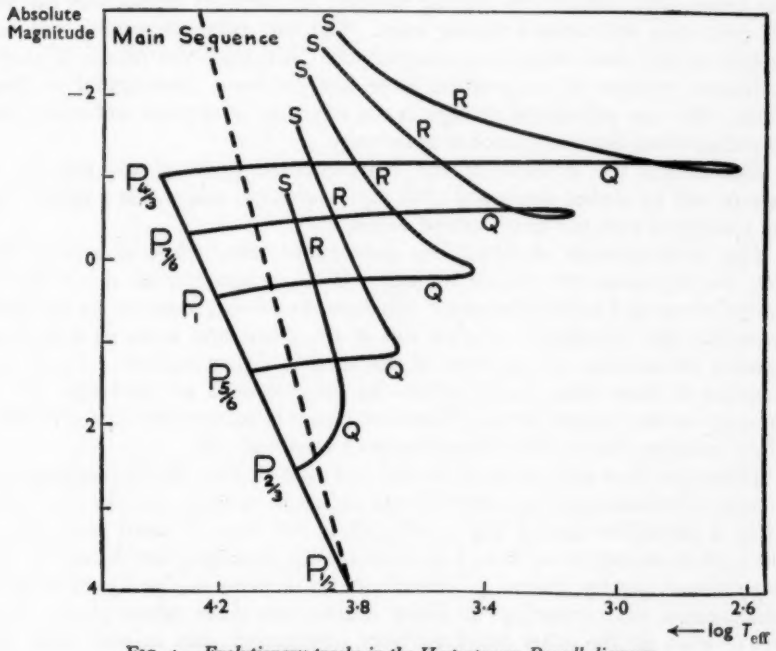


FIG. 4.—*Evolutionary tracks in the Hertzsprung–Russell diagram.*

$$\frac{\text{Accreted mass}}{\text{Original mass}} = \begin{cases} 0 \text{ at } P. \\ 0{\cdot}2 \text{ at } Q. \\ 1{\cdot}5 \text{ at } R. \end{cases}$$

Beyond S part of the outer region is convective.

As ν increases from zero, the representative point of the star will move rapidly to the right in the diagram, reaching the neighbourhood of the corresponding point Q when ν has reached a value of about 0·2. As ν increases further, the point remains in the same neighbourhood, reaching the point R

only when ν has reached a value of 1·5. A further increase of ν moves the representative point upwards and to the left until it reaches the point marked S, when a convective zone appears in the outer envelope not far from the discontinuity. Although the computations used in this paper no longer apply beyond this point, it may easily be seen that in fact the same type of development continues. The evolutionary tracks all tend asymptotically to the dotted line in the diagram which represents the locus of representative points of homogeneous hydrogen stars of differing masses (the main sequence).

8. The interpretation of Fig. 4 is now quite easy. Assuming a constant rate of accretion the representative point would move very rapidly from P to Q, and then very slowly in the region Q–R, and would finally move slowly towards S. There would hence be a very small density of stars in the P–Q region and a high density in the Q–R region.

If accretion ceases at some stage, mixing becomes again the predominant agent of evolution. Since however accretion was supposed to be a much more rapid process than mixing, the new developments will proceed much more slowly. This argument is reinforced by the consideration that a considerable time must elapse before the angular velocity has diffused throughout the star, and only then will serious mixing start. The first effect of mixing will be a blurring of the discontinuity of composition. Although the effects of such a continuous change of composition have not yet been investigated it seems unlikely that any substantial change in the structure of the star will occur until the composition again approaches uniformity.

Accordingly the position of the representative point of the star in the diagram will be almost stationary after rapid accretion ceases, for a period very long compared with the time scale of accretion.

The developments sketched out presuppose that, while accretion takes place, the accretion rate should be high. This in turn implies a low relative velocity of star and interstellar cloud, the density of which must not be too small. Hence the star is unlikely to move out of the cloud very soon, so that a total accretion of less than 20 per cent of the mass is rather unlikely. Since total accretions of more than (say) 3 times the original mass are probably not too common, we may expect the representative points of stars of the type considered to form mainly a broad band along the points marked QR.

Allowance must now be made for the fact that all stars that underwent such a course of evolution did not start off with the same masses. But the chief effect of this is merely to expand Fig. 4 vertically. For stars of small mass did not have time to convert more than a fraction of their hydrogen into helium so that the μ_i cannot greatly exceed $\frac{1}{2}$. Hence at the bottom of the figure only the developments corresponding to lower curves can have taken place. Very massive stars on the other hand are very uncommon, and in their case high helium contents are probably rather likely. Accordingly the upper parts of the figure are presumably sparsely populated, with development centring on the upper curves.

Fig. 4 refers, of course, only to stars in which electron scattering predominates, but a comparison with Paper II of this series suggests that no very large changes are likely to arise if photoelectric opacity is included. For massive red giants the low density in any case probably implies that electron scattering predominates.

Although further study and detailed comparisons are greatly required, it seems eminently likely at this stage that the Hoyle–Lyttleton model of red giants offers an adequate explanation of all the main features of that part of the Hertzsprung–Russell diagram lying to the right of the main sequence.

Newnham College,
Cambridge:
1951 *February* 19.

Trinity College,
Cambridge.

References

(1) C. M. Bondi, *M.N.*, **110**, 275, 1950.
(2) C. M. Bondi and H. Bondi, *M.N.*, **110**, 287, 1950.
(3) F. Hoyle and R. A. Lyttleton, *M.N.*, **102**, 218, 1942.
(4) F. Hoyle and R. A. Lyttleton, *M.N.*, **109**, 614, 1949.
(5) C. M. Bondi and H. Bondi, *M.N.*, **109**, 62, 1949.
(6) H. Bondi, *M.N.*, **110**, 601, 1950.
(7) P. A. Sweet, *M.N.*, **110**, 548, 1950.

ON THE FIGURE OF A PLANET WITH HOMOGENEOUS SHELL AND CORE

Harold Jeffreys

(Received 1951 May 28)

Summary

The equations of equilibrium of a two-layer rotating body, with uniform densities in the shell and core, are solved without use of the Radau approximation. It is found that this approximation loses accuracy appreciably when the central condensation is stronger than in the Earth. It remains true, however, that the surface ellipticity and the dynamical ellipticity are much more sensitive to the ratio C/Ma^2 than to the radius of the core.

Wiechert's theory of the figure of the Earth supposed a uniform core and a uniform shell, and such models are therefore usually known as of Wiechert's type. His main contribution, however, was to include terms of the second order in the ellipticities and to obtain quantitative results in some detail. A first-order theory had already been given by Thomson and Tait*, who gave the equations but no quantitative results. A still earlier one of S. Haughton† is quoted by E. J. Routh‡; some quantitative results had been found, the author reaching the conclusion, which was considered surprising, that the shell must be more than 748 miles thick.

Such discussions have had little attention since Radau's discovery that for the Earth, on a hydrostatic theory, if

$$K=\frac{5}{2}\frac{C}{Ma^2}, \quad m=\frac{\omega^2 a}{g}, \quad H=\frac{C-A}{C}, \tag{1}$$

then if m is known, knowledge of any one of K, H and the ellipticity e determines the other two within the uncertainty of observation. A model of the type considered contains three adjustable constants, restricted only by the need to give the right mean density; then effectively the theory of the figure of the Earth gives only one more relation and the problem remains formally insoluble until another datum is found. Radau's theory is a close approximation for a wide class of radial distributions of density. It is not universally valid, in fact the accuracy begins to fail if $K<0{\cdot}8$; and even with the correct value $K\doteqdot\frac{2}{3}$, there is loss of accuracy if the core radius is less than $2^{-1/2}$ of the outer radius.§ Thus the approximation might apparently cease to be good with the actual radius of the core found from seismology; but Bullen showed‖ that the variations of the density in the core and shell restored its validity.

* W. Thomson and P. G. Tait, *Treatise on Natural Philosophy*, Section 831, Cambridge, 1883.
† S. Haughton, *Trans. Roy. Irish Acad.*, **22**, 251, 1851.
‡ E. J. Routh, *Analytical Statics*, Vol. 2, p. 161, 1902.
§ H. Jeffreys, *M.N., Geophys. Suppl.*, **1**, 121–124, 1924.
‖ K. E. Bullen, *M.N., Geophys. Suppl.*, **3**, 401, 1936.

It remains interesting, in view of the possibilities of discontinuities of density in other planets and stars, to investigate the theory of Wiechert models without using the Radau approximation, since some bodies may lie outside its range of accuracy. For Jupiter and Saturn, the perturbations of the satellites together with the Radau approximation make K near $\frac{2}{3}$ and $\frac{1}{2}$ respectively.* Using the first-order theory† and taking the forms of the shell and core boundaries as

$$r=a(1+eY_2); \quad r=a\alpha(1+e'Y_2); \tag{2}$$

$$Y_2=\tfrac{1}{3}-\cos^2\theta, \tag{3}$$

and the densities as ρ, $\rho(1+\eta)$, we find, from the conditions that the outer surface and the core boundary are level surfaces,

$$-\tfrac{2}{3}e+\eta\alpha^3(-e+\tfrac{3}{5}\alpha^2e')+\tfrac{1}{2}(1+\eta\alpha^3)m=0, \tag{4}$$

$$\tfrac{3}{5}e-(1+\tfrac{2}{5}\eta)e'+\tfrac{1}{2}(1+\eta\alpha^3)m=0 \tag{5}$$

and

$$H=\frac{e+\eta\alpha^5e'}{1+\eta\alpha^5}; \quad K=\frac{1+\eta\alpha^5}{1+\eta\alpha^3}. \tag{6}$$

For the Earth K is nearly $\frac{5}{6}$, and we get, after some algebra,

$$\frac{e}{m}=1-\frac{\frac{2}{5}\alpha^2(1+2\alpha+3\alpha^2+\frac{3}{2}\alpha^3)}{2+4\alpha+4\frac{2}{5}\alpha^2+29\frac{4}{5}\alpha^3+55\frac{1}{5}\alpha^4+27\frac{3}{5}\alpha^5}. \tag{7}$$

The second term is small for all values of α from 0 to 1. This essentially confirms Radau's result from a completely different approach, but our interest is now in the amount of the departures. The large coefficients in the denominator are much reduced for smaller values of K, and e/m varies more with α for fixed K. We also have

$$e'/m=30\alpha^3(1+\alpha)^2/D, \tag{8}$$

where D is the denominator in (7). The calculation has been carried out for $K=\frac{1}{2}$, $\frac{2}{3}$, $\frac{5}{6}$ and 1. For $K=1$, the homogeneous case, $e=e'=H=1{\cdot}25m$ for all values of α. The results for other values of K are given in Table I in units of the fourth decimal.

TABLE I

	$K=\frac{1}{2}$			$K=\frac{2}{3}$			$K=\frac{5}{6}$		
α	e/m	e'/m	H/m	e/m	e'/m	H/m	e/m	e'/m	H/m
0·0	7142	0	7142	8333	0	8333	10^4	0	10^4
0·1	7112	35	7040	8306	61	8264	9980	146	9960
0·2	7022	265	6740	8225	456	8064	9926	1044	9852
0·3	6885	798	6283	8085	1334	7751	9860	2786	9720
0·4	6725	1608	5750	7989	2574	7473	9808	4759	9615
0·5	6578	2437	5198	7886	3880	7219	9776	6422	9553
0·6	6444	3475	4774	7810	5019	7025	9761	7637	9522
0·7	6337	4317	4396	7763	5965	6899	9755	8480	9510
0·8	(6254	5086	4178)	7726	6700	6814	9754	9060	9507
0·9	(6199	5663	3914)	(7705	7263	6763)	9754	9466	9509
1·0	(6154	6154	...)	(7692	7692	...)	(9756	9756	...)

The shell density is positive when $\alpha^2<K$; values not satisfying this condition are indicated by brackets. When $\alpha\to1$, with $K<1$, the shell reduces to a negative surface density.

* H. Jeffreys, *M.N.*, **84**, 535, 1924.

† H. Jeffreys, *The Earth* ("Figure of the Earth" chapter), Cambridge, 1924.

Direct use of the Radau approximation gives

	$K=\frac{1}{2}$	$\frac{2}{3}$	$\frac{5}{6}$
η_a	2·0625	1·25	0·5625
$\psi(\eta_a)$	0·9171	0·9792	0·9997
e/m	0·6154	0·7692	0·9756
H/m	0·3846	0·6731	0·9512

As the approximation, in its simplest form, takes $\psi(\eta)=1$, we should have expected that its results would be rough for $K=\frac{2}{3}$ and seriously wrong for $K=\frac{1}{2}$. Even for $K=\frac{5}{6}$, η is monotonic only if $\alpha>0{\cdot}707$. In the interval $0{\cdot}707<\alpha<0{\cdot}913(=K^{1/2})$, where the approximation would be expected to be good, it is in fact very close, but outside this interval it is not very good.

The main interest of this discussion rests in the possibility of determining two of the ratios C/Ma^2, e/m and H/m when the other is given; and we find that for $C/Ma^2>\frac{4}{15}$ a rough determination is possible, but not for smaller values.

An attempt has been made to see whether the present method by itself would lead to any useful determination of α when two of the ratios are given. From values found in my study* I find for the Earth

$$H/m=0{\cdot}9496\pm0{\cdot}0002;\quad e/m=0{\cdot}9760\pm0{\cdot}0020.$$

The latter depends on the data other than those derived from the constant of precession. Taking $K=\frac{5}{6}$, $\alpha=0{\cdot}5$ as trial values, I find a least squares solution

$$K=0{\cdot}8359\pm0{\cdot}0036;\quad \alpha=0{\cdot}72\pm0{\cdot}13.$$

K is nearly right, but α deviates from the true value by more than the standard error. This is of course attributable to neglect of second-order terms and of differences of density within the shell and core.

St. John's College,
Cambridge:
1951 *May* 26.

* H. Jeffreys, *M.N., Geophys. Suppl.*, **5**, 219–247, 1948.

THE SUN'S ROTATION DERIVED FROM SUNSPOTS 1934–1944 AND ADDITIONAL RESULTS

H. W. Newton and M. L. Nunn

(Communicated by the Astronomer Royal)

(Received 1951 May 26)

Summary

Results are given for the Sun's daily sidereal motion (ξ) derived from 136 single recurrent sunspots for the cycle 1934–1944. The expression relating ξ to solar latitude (ϕ) so derived, $\xi = 14^\circ{\cdot}38 - 2^\circ{\cdot}96 \sin^2\phi$, is in satisfactory accordance with values from similar data for the five previous 11-year cycles. A good internal accordance of measures is found by comparing ξ values derived from a full rotation period interval (27 days) with the mean (ξ_1), given by the two separate disk-passages of the same spot.

From the combined data of six sunspot cycles there appears no measurable variation in the Sun's rotation period (1) with alternate 11-year cycles, (2) within the (mean) cycle itself. Neither do the two hemispheres, north and south, behave differently.

Although not included in the discussion, the derived expression for the daily sidereal motion from non-recurrent spots, 1934–1944, is stated as a point of interest in the Addendum.

Introduction.—The problem of the Sun's rotation still deserves more observational data. Considerable differences in spectroscopic values cannot wholly be ascribed to differences of level in the chromosphere for the particular line or lines used. Moreover, St. John pointed out in 1932 * that collective spectroscopic results "show a progressive change in the equatorial rotation determined from the reversing layer during the past thirty years greater than can be attributed to errors of observation". The overall range in the (smoothed) observational values from 1906 to about 1925 is equivalent to no less than two days in the respective rotation periods for the Sun's equator.

Values derived from sunspots are more accordant *inter se*. The range in the equatorial value (a) for the daily sidereal rotation (ξ) at the equator from the expression $\xi = a - b\sin^2\phi$ given by some half-dozen investigations † is $0^\circ{\cdot}22$, equivalent to about 0·4 day in the rotation period.

Very consistent values over five sunspot cycles (1878–1933) have been given by the Greenwich measures of individual spots of long duration.‡ The range in the value of a does not exceed $0^\circ{\cdot}04$ per day. The dispersions of ξ values in the 5° zones of latitude, into which the data are divided, is also much less than in determinations based on spot *groups* (as in the investigation by Mr and Mrs Maunder §), especially those that are shorter-lived.

In the present paper, the results from recurrent spots for a sixth sunspot cycle (1934–1944) are presented. Over the whole period since 1878 there are now sufficient homogeneous data from at least two of the 5° latitude zones to

* C. E. St. John, *Trans. I.A.U.*, **4**, 43, 1933; see also *P.A.S.P.*, **47**, 295, 1935.

† The smallest value of a is $14^\circ{\cdot}35$ (Spoerer) and the largest $14^\circ{\cdot}57$ (O. A. Akesson).

‡ Royal Observatory, Greenwich, *M.N.*, **85**, 548, 1925. H. W. Newton, *M.N.*, **95**, 62, 1934.

§ E. W. and A. S. D. Maunder, *M.N.*, **65**, 813, 1905. O. A. Akesson, *Medd. Lunds Astronomiska Observatorium*, Serie II, Nr. 11, 1913.

show whether ξ varies within measurable limits (1) alternately from one cycle to the next, as do the magnetic polarities of sunspots, and (2) with the phase of the (mean) 11-year sunspot cycle. Results of this enquiry are stated herein.

Data.—The basic data comprise individual recurrent sunspots tabulated in Ledger I of the *Greenwich Photoheliographic Results* for the years 1934 to 1944. The volumes for 1937 and 1939 are in course of publication, those for 1940 to 1944 being still in MS.

Although the advantages in using stable recurrent spots appear far to outweigh the disadvantages, it is true nevertheless that the non-uniform characteristic motions*, especially of the leader spots of groups, are imposed on the true values of ξ. The present data include a large proportion of leader spots first seen as such or later as the sole survivor of the group, when they would be classified as "unipolar" spots.

Counting each appearance separately, the spots used in this paper were as follows:—

Leader spots (89): Followers (8): Other components (1):
Single spots, i.e. "unipolar" (174).

Any systematic effect from non-uniform proper motion on values of ξ is much reduced by (1) the data including spots in all stages of development, and (2) the use of the long time-interval of a solar rotation for deriving the values of ξ.

Regarding the stage of spot development, the distribution in three broad divisions for the present data is as follows:—

		Number	Mean area†
Phase I	Origin or early development to maximum	63	394
II	Middle life	139	298
III	Late decline (and extinction)	70	71

The time spent in each of the three phases is very approximately in the ratio 1 : 2 : 1. Phase II is sometimes markedly prolonged with a spot of area 200 or 300 millionths.

Derivation of ξ.—The method of reduction as used in earlier Greenwich papers is simply as follows:—

The mean longitude of the spot during its first complete (or part) disk-passage is compared with the corresponding mean longitude at its second appearance. Days on which any spot was $>80°$ from the central meridian are excluded.

If the change in longitude be $x°$ in the interval y mean solar days, then the diurnal angular motion is $x°/y$ relative to a meridian on the Sun having a sidereal rotation of 25·38 days‡ or $14°{\cdot}18_4$ per day, this being the datum adopted in the *Nautical Almanac* for computing the daily values of the longitude of the central meridian. Thus the daily angular sidereal motion, ξ, per mean solar day is $14°{\cdot}18+x°/y$. In rounding off mean values of ξ to two decimal places of a degree (as in Table I), the correction of $+0°{\cdot}004$, being the excess over $14°{\cdot}18$, has been taken into account.

Observational checks.—The comparison of the mean position of the same sunspot 27 or 28 days apart implies a (close) symmetrical distribution of the measures of longitude with respect to the central meridian. In the case of

* Royal Observatory, Greenwich, *M.N.*, **85**, 553, 1925.

† Areas, corrected for foreshortening, are expressed in millionths of the Sun's hemisphere.

‡ It is perhaps worth recording that this was the period adopted by Carrington because of "its admitting conveniently of much subdivision without remainders". He found later that this value applied to latitude 14°.

incomplete disk-passages of spots, the daily longitudes for comparison at the two epochs were selected so as to make the mean distances from the central meridian (and on the same side of it) correspond as nearly as possible. In this way, any small systematic displacements that may be present in the measures and which are a function of the distance from the centre of the disk will be eliminated. A small systematic error of this nature—zero at the centre of the Sun's disk and a maximum at the limb—could arise from differential foreshortening (not allowed for by the measurer) of the area subtended by the larger spots approaching the limb. This will tend to make the measured distance, r, of the spot from the disk centre smaller than its true value.

Other displacements have been suggested* as arising from (*a*) refraction above the photosphere producing an augmentation of the Sun's radius measured as R (irradiation will also produce an augmentation), and (*b*) a possible "depth-parallax" arising from a difference of level between spots and the effective limb. The combined effect of these possible displacements would tend to decrease the ratio r/R which enters directly into the computation of longitude and latitude of the spot, and thus decreases the derived daily angular motion. Although the use of recurrent spots avoids these systematic errors inherent in single disk-passage determinations, there is occasional uncertainty in the continuity of the longer-lived spots from one rotation to another. A general check is therefore desirable by comparing ξ values given by the same spot from the 27-day interval and the mean value (ξ_1) given by the separate disk-passages. Cases of complete or nearly complete disk-passages only have been included for this purpose. For each of the two disk-passages the mean longitude from the first five days is compared with that from the last five days. Days on which the spot position exceeded 80° from the central meridian are excluded, so that the reduced effect of the small possible displacements considered above would be unlikely to invalidate the check, as such, on the continuity of any particular sunspot.

The distribution of values, $\xi - \xi_1$, from 86 pairs of disk-passages is shown in Fig. 1, which also incorporates the development phases of the spots concerned. With respect to the centre grouping of values, say from $-0°{\cdot}10$ to $+0°{\cdot}10$ (comprising 63 of the 86 values), it may be said that the distribution, when slightly smoothed, is essentially that of a normal error distribution for a probable error of $\pm 0°{\cdot}03$ (single value) and displaced $+0°{\cdot}02$ from the origin.

This small positive asymmetry of $0°{\cdot}02$, which is five times the p.e. of the mean of the above 63 observations, may be real and represent the combined effects of the small systematic errors entering into a single disk-passage determination of the solar rotation.

It may be added that differential foreshortening alone would amount, on purely geometrical grounds, to about $0°{\cdot}01$ in the above derivation of ξ_1 values from circular spots of the average order of 3° in diameter or 350 millionths of the Sun's hemisphere in area.† A systematic error of this order could well arise from the optical distortion of the lens system of a photoheliograph, as instanced by the small corrections used since 1926 on this account for the Greenwich photographs.

* Papers of the International Union of Solar Research, Computing Bureau: R. S. Capon, *M.N.*, **73**, 361 and 732, 1913, in which various references are included.

† We are indebted to Mr H. F. Finch for the evaluation of this effect after an informal discussion on possible systematic errors.

The large negative values in Fig. 1 are found to be all* associated with Phase I of leader spots, when their forward motion in longitude would entirely obscure the other very small displacements.

A few other high values of $\xi - \xi_1$ indicated doubtful continuity of the spots on the Sun's invisible hemisphere, and this uncertainty was confirmed by a re-examination of their daily positions and areas. Such cases are not included in Fig. 1 nor in the data from which ξ was derived.

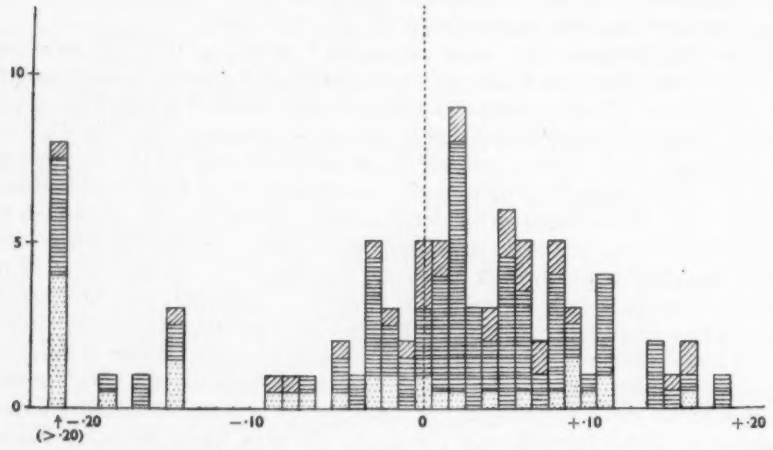


FIG. 1.—*Frequency distribution of $\xi - \xi_1$ values and phase development of the associated recurrent sunspots at both disk-passages.*

Ordinates: Frequency.

Abscissae: Values of $\xi - \xi_1$.

Phase I.

Phase II.

Phase III.

The overall mean, $\xi - \xi_1$, is less than 0°·01 (−0°·007). This small statistical difference not only substantiates the continuity of the spots provided by Ledger I, but suggests that non-recurrent spots from Ledger II can serve at least as a useful independent check for Table I.

The 282 non-recurrent spots selected include some from Ledger I which do not meet the requirements of a 27-day interval comparison. The derived values of the solar diurnal motion from these single disk-passages (designated ξ_0) are tabulated in the Addendum.

Results.—In Fig. 2 the separate values of ξ from 136 recurrent sunspots are plotted against their respective solar latitudes (ϕ). The same values averaged for 5° zones of latitude (0°·0 to 4°·9; 5°·0 to 9°·9, etc.) are given in Table I with northern and southern hemispheres shown separately in satisfactory accordance.

* Each recurrent spot is of course assigned two indices of phase development. Phase I scarcely ever extends to the second passage. Hence in Fig. 1, a 50 per cent occurrence of Phase I indicates a complete association of these spots with Phase I at their first disk-passage.

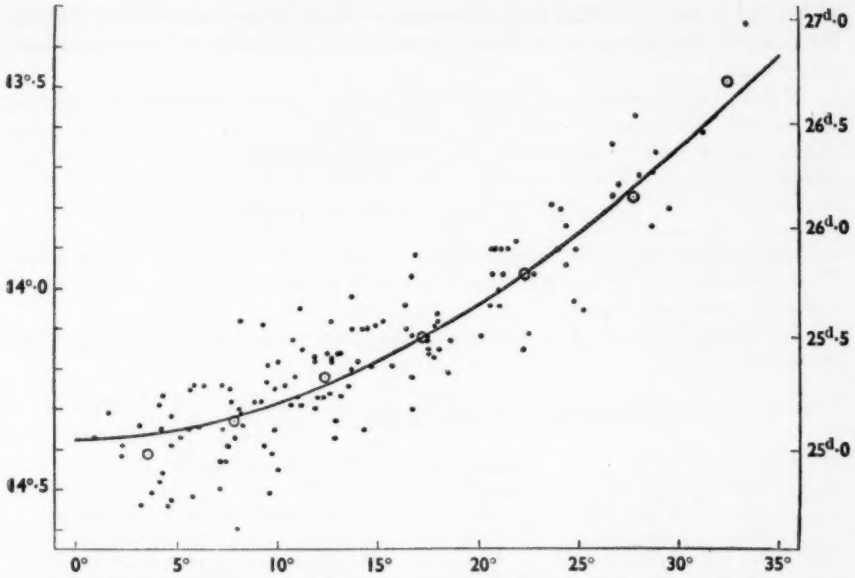


FIG. 2.—*Observed values of ξ from recurrent sunspots, 1934–1944. Means for each 5° zone of latitude are represented by open circles. The curve is computed from the expression*

$$\xi = 14°{\cdot}38 - 2°{\cdot}96 \sin^2 \phi.$$

Ordinates (left-hand side): Daily sidereal motion.
Ordinates (right-hand side): Sidereal rotation period in days.
Abscissae: Solar latitude.

TABLE I

Mean daily sidereal motions of sunspots, 1934–1944

		N. spots	S. spots	Mean	Mean dispersion	O–C
0°–5°	φ	3°·4	4°·0	3°·6		
	ξ	*14°·40*	*14°·43*	*14°·41 ± 0°·02*	±0°·08	+0°·04
	n	10	6	16		
5°–10°	φ	7°·7	7°·8	7°·8		
	ξ	*14°·32*	*14°·34*	*14°·33 ± 0°·01*	±0°·09	0°·00
	n	15	14	29		
10°–15°	φ	12°·9	11°·9	12°·4		
	ξ	*14°·20*	*14°·23*	*14°·22 ± 0°·01*	±0°·08	−0°·02₅
	n	18	18	36		
15°–20°	φ	17°·2	17°·2	17°·2		
	ξ	*14°·13*	*14°·12*	*14°·12 ± 0°·01*	±0°·06	0°·00
	n	7	15	22		
20°–25°	φ	22°·0	22°·7	22°·3		
	ξ	*14°·01*	*13°·90*	*13°·96 ± 0°·02*	±0°·08	+0°·01
	n	11	9	20		
25°–30°	φ	28°·4	27°·3	27°·6		
	ξ	*13°·76*	*13°·77*	*13°·77 ± 0°·03*	±0°·10	+0°·02
	n	3	8	11		
30°–35°	φ	32°·4		32°·4		
	ξ	*13°·48*	...	*13°·48*	±0°·14	−0°·05
	n	2		2		

In column 5 the mean dispersion (or scatter of observed points in Fig. 1) is given, from which the respective probable errors in the preceding column have been computed.

Taking the observed mean values of ξ in each 5° belt of latitude and weighting the values according to the respective number of observations (the mean dispersion in each 5° belt being comparable), a solution by least squares was obtained, giving

$$\xi = 14^\circ{\cdot}38_2 \pm 0^\circ{\cdot}009 - (2^\circ{\cdot}96 \pm 0^\circ{\cdot}09)\sin^2\phi.$$

Residuals (observed *minus* computed) are given in the last column of the table. This evaluation of ξ is in good accordance with those for the five earlier sunspot cycles as shown in Table II.

In Table II this result is compared with those yielded by the preceding five cycles.*

TABLE II

Solutions for ξ (recurrent sunspot data) over six cycles

Cycle	Years	Number of sunspots	Diurnal motion (ξ)
I	1878–1888	89	$14^\circ{\cdot}34 \pm 0^\circ{\cdot}01 - (2^\circ{\cdot}35 \pm 0^\circ{\cdot}08)\sin^2\phi$
II	1889–1899	122	$14^\circ{\cdot}36 \pm 0^\circ{\cdot}01 - (2^\circ{\cdot}58 \pm 0^\circ{\cdot}18)\sin^2\phi$
III	1900–1913	99	$14^\circ{\cdot}38 \pm 0^\circ{\cdot}01 - (2^\circ{\cdot}69 \pm 0^\circ{\cdot}13)\sin^2\phi$
IV	1914–1923	139	$14^\circ{\cdot}38 \pm 0^\circ{\cdot}01 - (2^\circ{\cdot}50 \pm 0^\circ{\cdot}18)\sin^2\phi$
V	1924–1933	130	$14^\circ{\cdot}37 \pm 0^\circ{\cdot}01 - (3^\circ{\cdot}01 \pm 0^\circ{\cdot}11)\sin^2\phi$
VI	1934–1944	136	$14^\circ{\cdot}38 \pm 0^\circ{\cdot}01 - (2^\circ{\cdot}96 \pm 0^\circ{\cdot}09)\sin^2\phi$

Using the present solution (6th cycle), values of ξ at 5° steps of latitude are given in Table III, together with the equivalent rotation periods. These values will be found to compare best with those derived by Carrington from his visual measures of sunspots 1853 to 1861.

TABLE III

Computed values of ξ and equivalent rotation periods from $\xi = 14^\circ{\cdot}38_2 - 2^\circ{\cdot}96\sin^2\phi$

Solar latitude	Daily sidereal motion, ξ	Sidereal period	Mean synodic period
	°	d	d
0°	14·38	25·03	26·87
5°	14·36	25·07	26·92
10°	14·29	25·19	27·06
15°	14·18	25·38	27·28
20°	14·04	25·65	27·59
25°	13·85	25·99	27·98
30°	13·64	26·39	28·45
35°	13·41	26·85	28·98

Test for periodic variation of ξ.—(1) Alternation with cycle; (2) during mean cycle.

(1) *Alternation with cycle.*—The mean values of ξ for the odd and even cycles have been taken from Table II and weighted according to the respective relative

* These solutions differ slightly from those given in *M.N.*, **95**, 62, 1934 (Table II) for which a less rigorous solution was used. In that table, lines 1 and 2, *read* 14°·35 and 14°·38 respectively for 14°·36 and 14°·39.

weights of a and b in each of the six solutions. The derived expressions, in close accordance, are as follows:—

Cycles I, III, V. $\xi = 14^\circ{\cdot}36_4 \pm 0^\circ{\cdot}005 - (2^\circ{\cdot}60 \pm 0^\circ{\cdot}06)\sin^2\phi$,

Cycles II, IV, VI. $\xi = 14^\circ{\cdot}37_5 \pm 0^\circ{\cdot}007 - (2^\circ{\cdot}82 \pm 0^\circ{\cdot}07)\sin^2\phi$.

However, the derivation of the equatorial value from an expression which best fits the whole range of latitude values (with considerable weight because of numbers in the middle latitudes) might obscure some small significant differences near the equator itself. We therefore give the observed values of ξ for the equatorial zone, 0°–5°, and also these values reduced to the equator by the present solution for ξ.

	Observed ξ	ϕ	ξ at 0°
Odd cycles	$14^\circ{\cdot}36_0 \pm 0^\circ{\cdot}01_5$	3°·1	*14°·36₉*
Even cycles	$14^\circ{\cdot}40_9 \pm 0^\circ{\cdot}01_1$	2°·9	*14°·41₇*

The difference of 0°·05, being less than three times its probable error, is therefore of very doubtful significance.

The two most prolific 5° zones of latitude (10°–15°; 15°–20°) provide the following observed mean values over the same extended period, 1878–1944. Values of ξ in italics indicate the slightly adjusted values to a common mean latitude of 12°·5 for the lower zone and to 17°·2 for the higher. n is the number of spots.

TABLE IV

Cycles	Zone 10°–15°			Zone 15°–20°		
	ξ	ϕ	n	ξ	ϕ	n
Odd	$14^\circ{\cdot}23_6$	12°·7	86	$14^\circ{\cdot}13_1$	17°·0	73
	14°·24₀			*14°·12₆*		
Even	$14^\circ{\cdot}23_0$	12°·3	115	$14^\circ{\cdot}12_2$	17°·5	73
	14°·22₇			*14°·13₀*		

From these values we see that there is no observational evidence of any cyclical variation of ξ in these latitude ranges as great as 1 part in 1000 in a 22-year period.

(2) *Variation of ξ within the 11-year cycle.*—To test whether there is any periodic variation of ξ during the (mean) sunspot cycle, the data for the above two zones were divided into four epochs according to time-interval from solar maximum. The epochs of the six maxima were adopted from the Zürich data of sunspot "numbers", and the limits of each of the four epochs within the mean 11-year cycle were adjusted so as to include as nearly as possible the same number of observations in the same 5° zone. The results are set forth in Table V. For more exact comparison, values of ξ in italics include the small correction necessary to bring the values appropriate to the mean latitude of 12°·5 in the case of the first zone and of 17°·2 for the other zone. Although the greatest difference between one epoch and the next is 0°·08 (i.e. about four times the probable error), the overall run of values suggests no real variation with phase of the cycle.

This result from sunspots (with the reservation that the solar minimum itself is barely represented) contrasts sharply with those typical of the upper chromosphere for which Evershed*, for instance, obtained from measures of the H and K lines in prominence spectra, a difference of 2° *per day* as between solar minimum and times of greatest activity.

* J. Evershed, *M.N.*, **105**, 205, 1945.

TABLE V

ξ values during the (mean) sunspot cycle

Zone		Epoch from sunspot maximum (=0^y·0)			
		1	2	3	4
10°–15°	*e*	−1^y·7	−0^y·1	+1^y·3	+3^y·6
	φ	12°·6	12°·5	12°·5	12°·3
	ξ	14°·22	14°·26	14°·22	14°·23
		14°·22	*14°·26*	*14°·22*	*14°·23*
	n	52	51	50	48
15°–20°	*e*	−2^y·3	−1^y·0	+0^y·5	+2^y·5
	φ	17°·7	17°·8	16°·8	16°·6
	ξ	14°·17	14°·11	14°·11	14°·12
		14°·18	*14°·13*	*14°·10*	*14°·10*
	n	34	37	37	38

Mean *ξ*, 1878–1944: (10°–15°) 14°·233 at latitude 12°·5 (201 spots).
(15°–20°) 14°·126 at latitude 17°·2 (146 spots).

Northern and southern hemispheres compared (1878–1944).—Finally from the combined data, a solution for *ξ* has been made for the separate hemispheres * to see whether there is any measurable difference in behaviour. It will be seen that there is no difference within the range of present accuracy.

North: $\xi = 14°{\cdot}37_2 \pm 0°{\cdot}010 - (2°{\cdot}68 \pm 0°{\cdot}12)\sin^2\phi$ (357 spots),
South: $\xi = 14°{\cdot}38_5 \pm 0°{\cdot}005 - (2°{\cdot}91 \pm 0°{\cdot}07)\sin^2\phi$ (358 spots).

Combining data for six cycles by 5° latitude zones and re-solving, we obtain

$$\xi = 14°{\cdot}37_7 \pm 0°{\cdot}006 - (2°{\cdot}77 \pm 0°{\cdot}08)\sin^2\phi.$$

It is of interest to compare the above solution with that obtained from the six individual solutions in Table II combined according to the weights of *a* and *b*, as follows:—

$$\xi = 14°{\cdot}36_8 \pm 0°{\cdot}004 - (2°{\cdot}69 \pm 0°{\cdot}04)\sin^2\phi.$$

Acknowledgment is gratefully made to the Astronomer Royal for the opportunity given to prepare this paper; for advice concerning the weighted solution adopted, and also for his suggestion that the data might be arranged to show any existing periodic variation of *ξ*.

With the kind cooperation of the Superintendent, H.M. Nautical Almanac Office, the various data on the original working cards were transferred to punched cards and so passed through the Hollerith machine to give the collected values in Table I and in the Addendum, together with other information.

Addendum

Using data from the non-recurrent sunspots for the period 1934–1944, values (designated ξ_0) of the daily sidereal motion have been derived from 5-day means of longitude as for ξ_1 values (i. e. separate disk-passages of recurrent spots).

Of the 282 spots available, mainly from Ledger II of the *Greenwich Photoheliographic Results* but including a few components of groups in Ledger I that were not suitable for the basic discussion, 89 spots were seen as leaders of groups,

* As for the first four cycles in *M.N.*, **85**, 552, 1925.

41 as followers, three as other components and 149 as single or "unipolar" spots. The distribution of their development-phases is as follows, the mean areas being added in brackets:—

Phase I—20 (309); Phase II—170 (236); Phase III—92 (122).

The above phase assessment is necessarily more approximate than for the long-lived spots. Sometimes, the whole life-history of the spot was completed during its disk-passage, and in such cases the index II was assigned.

Mean values of ξ_0 in 5° belts of latitude were derived as for ξ values in Table I, and a solution made as before, which gave:—

$$\xi = 14^\circ{\cdot}41_4 \pm 0^\circ{\cdot}004 - (3^\circ{\cdot}36 \pm 0^\circ{\cdot}04)\sin^2\phi.$$

The residuals (O–C) for the respective 5° latitude belts are satisfactorily small and read as follows from the equatorial belt to that for 30°–35°, the respective numbers of spots being added in brackets:—

$+0^\circ{\cdot}02$ (9); $+0^\circ{\cdot}01$ (58); $-0^\circ{\cdot}01$ (87); $0^\circ{\cdot}00$ (54); $-0^\circ{\cdot}01$ (41); $+0^\circ{\cdot}02$ (20); and $0^\circ{\cdot}00$ (13).

It seems worth recording the above subsidiary result without stressing the possible significance of the larger values of the a and b coefficients as compared with those derived from long-lived spots. Although a small physical difference (e. g. level difference) between these two classifications of spots may be indicated, it is nevertheless true that the values derived for the solar rotation depend to some extent on the selection of spots used and their phase and rate of development. A more rapid development of the shorter-lived spots may contribute larger values of ξ_0 (contrary to the effect of small systematic errors previously mentioned on p. 415), since the forward movements in longitude of leaders during growth will be effectively averaged over a 9 or 10 day interval instead of 27 days, as for the recurrent spots.

Royal Greenwich Observatory,
Herstmonceux Castle,
Sussex:
1951 *May* 19.

MEAN AREAS AND HELIOGRAPHIC LATITUDES OF SUNSPOTS IN THE YEAR 1946

Royal Observatory, Greenwich

(Communicated by the Astronomer Royal)

(Received 1951 March 22)

The following results are in continuation of those given in *M.N.*, **110,** 501, 1950, and are derived from the measurement at Greenwich and Herstmonceux of photographs taken at the Royal Observatories of Greenwich and the Cape and at the Kodaikanal Observatory, India.

Table I gives the mean daily areas of umbrae, whole spots and faculae for each synodic rotation of the Sun included in the year 1946; the means for each year as a whole are included in Table II, which summarizes the yearly values since the last sunspot minimum of 1944.

TABLE I

No. of Rotation	Rotation Commenced U.T. 1946	Days Photographed	Mean Daily Areas: Projected* Umbrae	Projected* Whole Spots	Projected* Faculae	Corrected for Foreshortening† Umbrae	Corrected for Foreshortening† Whole Spots	Corrected for Foreshortening† Faculae
1235	Jan. 4·63	27	295	1648	1093	242	1376	1353
1236	Jan. 31·97	27	709	3886	1619	550	3108	1917
1237	Feb. 28·31	28	306	1882	1846	255	1600	2210
1238	Mar. 27·62	27	296	1695	1826	240	1401	2201
1239	Apr. 23·89	27	299	1806	1699	210	1282	2074
1240	May 21·12	27	149	854	1653	111	649	2049
1241	June 17·32	28	247	1526	1844	184	1143	2285
1242	July 14·52	27	727	5170	1905	537	3915	2427
1243	Aug. 10·74	27	439	2534	1600	311	1836	1966
1244	Sept. 6·98	27	377	2351	1799	276	1732	2185
1245	Oct. 4·26	28	328	1924	2065	236	1387	2552
1246	Oct. 31·55	27	418	2629	2046	295	1896	2489
1247	Nov. 27·86	27	489	3167	2338	359	2392	2815

* Expressed in millionths of the Sun's disk.

† Expressed in millionths of the Sun's hemisphere.

TABLE II

Mean Areas and Heliographic Latitudes

Year	No. of Days: Photographed	No. of Days: Without Spots	Mean Daily Areas: Projected* Umbrae	Projected* Whole Spots	Projected* Faculae	Corrected for Foreshortening† Umbrae	Corrected for Foreshortening† Whole Spots	Corrected for Foreshortening† Faculae
1944	366	157	30	160	281	23	126	338
1945	365	14	102	560	773	78	429	939
1946	365	0	389	2381	1793	291	1817	2188

* Expressed in millionths of the Sun's disk.

† Expressed in millionths of the Sun's hemisphere.

Table III gives for each rotation in the year 1946 the mean daily area of the whole spots (corrected for foreshortening) and the mean heliographic latitude of the spotted areas for both northern and southern hemispheres.

The mean heliographic latitude of the entire spotted area and the mean distance from the equator of all spots are also tabulated. The mean values for the year 1946 are included in Table IV.

Tables II and IV are in continuation of similar tables in *Monthly Notices*: for the years *1874* to *1888*, **49,** 381, 1889: *1889* to *1902*, **63,** 465, 1903: *1901* to *1914*, **76,** 402, 1916: *1913* to *1924*, **85,** 1007, 1925: *1923* to *1933*, **94,** 870, 1934; and *1933* to *1945*, **110,** 501, 1950.

The rotations in Tables I and III are numbered in continuation of Carrington's series.

The chief features of the record are as follows:—

(1) The increasing steepness in the rise in sunspot activity after a solar minimum in 1944·2. As compared with the year 1945, the mean daily sunspot area has risen more than four times and the faculae $2\frac{1}{4}$ times (Table II).

The large mean daily sunspot area value of 3108 millionths (corrected area) of rotation 1236 (Table I) is largely due to the giant sunspot which, although supplanted in 1947, was up to this date the largest ever recorded at Greenwich. Similarly, a second giant spot in July contributed largely to the high value of 3915 in rotation 1242.

The run of the mean daily areas of faculae from rotation to rotation is smoother, because of the much longer life of faculae arising with spots.

TABLE III

No. of Rotation	Rotation Commenced U.T.	Spots North of the Equator		Spots South of the Equator		Mean Latitude of Entire Spotted Area	Mean Distance from Equator of all Spots
		Mean Daily Area	Mean Heliographic Latitude	Mean Daily Area	Mean Heliographic Latitude		
	1946		°		°	°	°
1235	Jan. 4·63	905	22·15	471	16·29	+ 8·99	20·14
1236	Jan. 31·97	2514	25·44	594	19·98	+16·77	24·40
1237	Feb. 28·31	943	25·72	657	22·73	+ 5·81	24·49
1238	Mar. 27·62	510	26·60	891	30·48	− 9·71	29·07
1239	Apr. 23·89	487	24·58	796	21·54	− 4·03	22·69
1240	May 21·12	510	19·83	139	20·30	+11·26	19·93
1241	June 17·32	522	24·10	620	22·81	− 1·37	23·40
1242	July 14·52	2721	20·24	1194	21·89	+ 7·39	20·74
1243	Aug. 10·74	1434	18·04	401	16·70	+10·45	17·75
1244	Sept. 6·98	1295	18·17	436	15·04	+ 9·81	17·38
1245	Oct. 4·26	519	15·69	868	14·83	− 3·41	15·15
1246	Oct. 31·55	1192	15·27	704	13·31	+ 4·66	14·54
1247	Nov. 27·86	1052	14·93	1341	10·31	+ 0·78	12·34

(2) On no day in 1946 were sunspots and faculae absent from the Sun's disk.

(3) The mean daily area of the northern hemisphere was greatly in excess of that of the southern. But of the 437 units of the northern excess, 309 units were alone due to the two giant spots mentioned above. In actual numbers of groups, the southern was rather the more prolific, as shown next.

TABLE IV

Year	Days Photographed	Spots North of the Equator: Mean Daily Area	Spots North of the Equator: Mean Heliographic Latitude	Spots South of the Equator: Mean Daily Area	Spots South of the Equator: Mean Heliographic Latitude	Mean Latitude of Entire Spotted Area	Mean Distance from Equator of all Spots
			°		°	°	°
1944	366	42	19·00	83	22·81	−8·70	21·53
1944	366 Old Cycle	7	4·17	6	7·61	−1·14	5·72
	366 New Cycle	35	22·18	77	24·02	−9·63	23·45
1945	365	121	20·13	309	20·26	−8·92	20·22
1946	365	1127	20·74	690	18·79	+5·73	20·00

(4) The number and distribution, northern and southern hemisphere, of spot groups of (*a*) two days or longer; (*b*) one day only, are as follows:—

	(*a*)	(*b*)
Northern spots	179	32
Southern spots	190	43
Total	369	75

The range in latitude for the longer-lived spots was from 6° to 42°. The latitudes of the 1-day spots ranged from 1° to 41°.

(5) The mean weighted latitude of all spots was 20°·0—practically the same as for 1945. The latitudes of the two giant spots were +26°·1 and +22°·2 respectively.

A short descriptive survey of the sunspots of 1946 has been given in *Monthly Notices*, **107**, 103, 1947.

An appended table gives the mean daily areas of sunspots and of faculae (corrected for foreshortening and expressed in millionths of the Sun's hemisphere) for each calendar month of 1946.

Monthly Mean Daily Areas of Sunspots and Faculae
1946

Month	Spots	Faculae	Month	Spots	Faculae
Jan.	1237	1285	July	2912	2431
Feb.	3091	1966	Aug.	2291	2010
Mar.	1551	2283	Sept.	1860	2181
Apr.	1433	2098	Oct.	1345	2564
May	1085	2036	Nov.	1852	2458
June	716	2099	Dec.	2514	2819

Royal Greenwich Observatory,
Herstmonceux Castle, Sussex:
1951 *March.*

CONTENTS

PAGE

Meeting of 1951 April 13:
- Fellows elected ... 345
- Junior Members elected ... 345
- Presents announced... ... 345
- Associates elected ... 345

Presentation of the Gold Medal ... 346

Meeting of 1951 May 11:
- Fellows elected ... 346
- Presents announced... ... 346

Antonie Pannekoek, The origin of astronomy (George Darwin Lecture) 347

R. Hanbury Brown and C. Hazard, Radio emission from the Andromeda nebula 357

M. J. Seaton, The chemical composition of the interstellar gas 368

J. B. Sykes, Approximate integration of the equation of transfer 377

W. S. Finsen, An eyepiece interferometer 387

H. E. Butler, Photoelectric recording of stellar occultations 393

C. M. Bondi and H. Bondi, Models for red giant stars. III. Models with a single discontinuity of composition and opacity due to electron scattering 397

Harold Jeffreys, On the figure of a planet with homogeneous shell and core ... 410

H. W. Newton and M. L. Nunn, The Sun's rotation derived from sunspots 1934–1944 and additional results 413

Royal Observatory, Greenwich, Mean areas and heliographic latitudes of sunspots in the year 1946 422

Printed in London by Taylor & Francis, Ltd., 1951 *December.*